Lecture Notes in Electrical Engineering 1182

Triwiyanto Triwiyanto
Achmad Rizal
Wahyu Caesarendra *Editors*

Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics

ICEBEHI 2023, 4–5 October, Surabaya, Indonesia

Springer

Lecture Notes in Electrical Engineering

Volume 1182

Series Editors

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Napoli, Italy
Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico
Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, München, Germany
Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China
Shanben Chen, School of Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China
Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore
Rüdiger Dillmann, University of Karlsruhe (TH) IAIM, Karlsruhe, Baden-Württemberg, Germany
Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China
Gianluigi Ferrari, Dipartimento di Ingegneria dell'Informazione, Sede Scientifica Università degli Studi di Parma, Parma, Italy
Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain
Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA
Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China
Janusz Kacprzyk, Intelligent Systems Laboratory, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
Alaa Khamis, Department of Mechatronics Engineering, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt
Torsten Kroeger, Intrinsic Innovation, Mountain View, CA, USA
Yong Li, College of Electrical and Information Engineering, Hunan University, Changsha, Hunan, China
Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA
Ferran Martín, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain
Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore
Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany
Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA
Subhas Mukhopadhyay, School of Engineering, Macquarie University, Sydney, NSW, Australia
Cun-Zheng Ning, Department of Electrical Engineering, Arizona State University, Tempe, AZ, USA
Toyoaki Nishida, Department of Intelligence Science and Technology, Kyoto University, Kyoto, Japan
Luca Oneto, Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genova, Genova, Genova, Italy
Bijaya Ketan Panigrahi, Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India
Federica Pascucci, Department di Ingegneria, Università degli Studi Roma Tre, Roma, Italy
Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China
Gan Woon Seng, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore
Joachim Speidel, Institute of Telecommunications, University of Stuttgart, Stuttgart, Germany
Germano Veiga, FEUP Campus, INESC Porto, Porto, Portugal
Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Haidian District Beijing, China
Walter Zamboni, Department of Computer Engineering, Electrical Engineering and Applied Mathematics, DIEM—Università degli studi di Salerno, Fisciano, Salerno, Italy
Kay Chen Tan, Department of Computing, Hong Kong Polytechnic University, Kowloon Tong, Hong Kong

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering—quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

Triwiyanto Triwiyanto · Achmad Rizal ·
Wahyu Caesarendra
Editors

Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics

ICEBEHI 2023, 4–5 October, Surabaya, Indonesia

Springer

Editors
Triwiyanto Triwiyanto
Department of Medical Electronics Technology
Poltekkes Kemenkes Surabaya
Surabaya, Indonesia

Achmad Rizal
School of Electrical Engineering
Telkom University
Bandung, Indonesia

Wahyu Caesarendra
Faculty of Integrated Technologies
Universiti Brunei Darussalam
Gadong, Brunei Darussalam

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-97-1462-9 ISBN 978-981-97-1463-6 (eBook)
https://doi.org/10.1007/978-981-97-1463-6

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Paper in this product is recyclable.

Organization

Organizing Committee

Dr. Endro Yulianto, Poltekkes Kemenkes Surabaya (Chairman)
Dr. Triwiyanto, Poltekkes Kemenkes Surabaya (Technical Program Committee Chair)
Sari Luthfiyah, Poltekkes Kemenke Surabaya (Secretary I)
Anita Miftahul Maghfiroh, Poltekkes Kemenkes Surabaya (Secretary II)
Triana Rahmwati, Poltekkes Kemenkes Surabaya
M. Prastawa Assalim T. P., Poltekkes Kemenkes Surabaya
Ridha Mak'ruf, Poltekkes Kemenkes Surabaya
Priyambada Cahya Nugraha, Poltekkes Kemenkes Surabaya
Farid Amrinsani, Poltekkes Kemenkes Surabaya
Syaifudin, Poltekkes Kemenkes Surabaya
Syevana Dita Musvika, Poltekkes Kemenkes Surabaya
Levana Forra Wakidi, Poltekkes Kemenkes Surabaya
Singgih Yudha Setiawan, Poltekkes Kemenkes Surabaya
Lusiana, Poltekkes Kemenkes Surabaya
Deenda Putri Duta Natalia, Poltekkes Kemenkes Surabaya

Proceeding Editorial Board

Dr. Triwiyanto, Poltekkes Kemenkes Surabaya, Indonesia
Dr. Achmad Rizal, Telkom University, Bandung, Indonesia
Wahyu Caesarendra, Ph.D., Universiti Brunei Darussalam, Brunei Darussalam

Scientific Program Committee

Achmad Rizal, Telkom University, Indonesia
Agung Triayudi, Universitas Nasional, Indonesia
Anggara Trisna Nugraha, Politeknik Perkapalan Negeri Surabaya, Indonesia
Alfin Hikmaturokhman, IT Telkom Purwokerto, Indonesia
Alfian Pramudita Putra, Biomedical Engineering, Faculty Science and Technology, Universitas Airlangga, Indonesia
Alfian Ma'arif, Universitas Ahmad Dahlan, Indonesia
Anilkumar Suthar, Rajasthan Technical University, India
Bambang Guruh Irianto, Poltekkes Kemenkes Surabaya, Indonesia
Candra Zonyfar, Universitas Buana Perjuangan Karawang, Indonesia
Devi Handaya, Politeknik Negeri Jakarta, Indonesia
Dwi Ely Kurniawan, Politeknik Negeri Batam, Indonesia
Deyby Huamanchahua Canchanya, Tecnológico de Monterrey, Mexico
Eka Legya Frannita, Universitas Gadjah Mada, Indonesia
Eko Supriyanto,Universiti Teknology Malaysia, Malaysia
Ejay Nsugbe,Cranfield University, Cranfield, United Kingdom
Faheem Reegu, Jazan University, Arab Saudi
Haresh Pandya, Saurashtra University, India
Hasanain Al-Behadili, University of Misan, Iraq
Harikrishna Parikh, Saurashtra University, India
Henderi, Universitas Raharja, Indonesia
Jasten Keneth D. Treceñe, Eastern Visayas State University, Tanauan Campus, Philippines
Johan Reimon Batmetan, Universitas Negeri Manado, Indonesia
Joseph Bamidele Awotunde, University of Ilorin, Ilorin, Kwara State, Nigeria
Kharudin Ali, UC TATI, Malaysia
Mas Aly Afandi, Institut Teknologi Telkom Purwokerto, Indonesia
Manas Kumar Yogi, Pragati Engineering College (Autonomous), India
Mera Kartika Delimayanti, Politeknik Negeri Jakarta, Indonesia
Michael G. Albino, President Ramon Magsaysay State University, Philippines
Nada Fitrieyatul Hikmah, Institut Teknologi Sepuluh Nopember, Indonesia
Phuoc-Hai Huynh, An Giang University, Vietnam
Pradeep N., Bapuji Institute of Engineering and Technology, Davangere, India
Raja Siti Nur Adiimah Raja Aris, UC TATI, Malaysia
Ramen A Purba, Politeknik Unggul LP3M, Medan, Indonesia
Rendra Gustriansyah, Universitas Indo Global Mandiri, Indonesia
Riky Tri Yunardi, Universitas Airlangga, Indonesia
Rismayani, STMIK Dipanegara Makassar, Indonesia
Sidharth Pancholi, Indian Institute of Technology Delhi, India
Shajedul Islam, Health Sciences University of Hokkaido, Japan
Suzanna Ridzuan Aw,University College TATI, Malaysia
Vishwajeet Shankhwar, National Institute of Technology Jalandhar, India

Vijay Anant Athavale, Walchand Institut of Technology, India, India
Wahyu Pamungkas, Institut Teknologi Telkom Purwokerto, Indonesisa
Wan Faizura Wan Tarmizi, UC TATI, Malaysia
Yohanssen Pratama, Institut Teknologi Del, Indonesia
Yuant Tiandho, Universitas Bangka Belitung, Indonesia

Preface

4th ICEBEHI-2023, the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics, takes place during October 4–5, 2023 on virtual platforms (Zoom app). The conference was organized by the Department of Medical Electronics Technology, Health Polytechnic Ministry of Health Surabaya (Poltekkes Kemenkes Surabaya), Surabaya, Indonesia, and Co-organized by Telkom University, Bandung, Indonesia, University College TATI, Malaysia, Universitas Trisakti, Indonesia, and Walchand Institute of Technology, India. With the aim of bringing together as a family all leading scientists, academicians, educationists, young scientists, research scholars, and students to present, discuss, and exchange their experiences, Innovation ideas, and recent development in the field of electronics, biomedical engineering, and health informatics.

The ICEBEHI-2023 conference was held online through Zoom application. More than 100 participants (presenter and non-presenter) attended the conference, they were from Vietnam, Malaysia, Russia, Brunei Darussalam, Azerbaijan, China, South Korea, Australia, Poland, Libya, India, and Indonesia. The scientific program of this conference included many topics related to electronics and biomedical engineering as well as those in related fields. In this conference, three distinguished keynote speakers and one invited speaker had delivered their research works in the area of Biomedical Engineering. Each keynote speech lasted 50 minutes.

ICEBEHI-2023 Conference collects the latest research results and applications on electronics, biomedical engineering, and health informatics. It includes a selection of 46 papers from 93 papers submitted to the conference from universities all over the world. All of the accepted papers were subjected to strict peer-reviewing by 2–3 expert referees. All articles have gone through a plagiarism check. The papers have been selected for this volume because of quality and relevance to the conference.

We are very grateful to the committee which contributed to the success of this conference. Also, we are thankful to the authors who submitted the papers, it was the quality of their presentations and communication with the other participants that really made this web conference fruitful. Last but not least, we are thankful to the Lecture Note in Electrical Engineering (Springer) Publishing for their support, it was not only the support but also an inspiration for organizers. We hope this conference

can be held every year to make it an ideal platform for people to share views and experiences in electronics, biomedical engineering, health informatics, and related areas. We are expecting you and more experts and scholars around the globe to join this international event next year.

Surabaya, Indonesia

Dr. Triwiyanto Triwiyanto
triwi@poltekkesdepkes-sby.ac.id

Contents

Multi Socket Transmission System Application with Advanced Encryption Standard Algorithm to Support Confidential Medical Data Security 1
Bita Parga Zen, Abdurahman, Anggi Zafia, Annisaa Utami, Iwan Nofi Yono Putro, and Faisal Dharma Aditama

The Application of Virtual Reality Using Kinect Sensor in Biomedical and Healthcare Environment: A Review 15
Henry Candra, Umi Yuniati, and Rifai Chai

Automated Taekwondo Kick Classification Using SVM and IMU Sensor on Arduino Nano 33 BLE 39
Qoriina Dwi Amalia, Azhar Agustian Gunawan, Grachia Salsabila Yulian, Achmad Rizal, and Istiqomah

Classification EEG Signal Using Texture Analysis and Artificial Neural Network for Alcoholic Detection 53
Donny Setiawan Beu, Hilal Hamdi Simatupang, Achmad Rizal, Rita Purnamasari, and Yunendah Nur Fuadah

CC1101 Network for Healthcare Cyber Physical System on Air Quality Data Acquisition 63
Firmansyah Maulana Sugiartana Nursuwars, Nurul Hiron, Aldy Putra Aldya, and Angga Setiawan Wahyudin

Modelling of Solar Irradiance for Optimal Solar-Powered Car Performance at EPIC Solar Farm Pathway 79
Afidatul Nadia Mok Hat, Ruzlaini Ghoni, Mohd Tarmizi Ibrahim, Ahmad Firdaus Zali, and Fuaad Mohamed Nawawi

Revolutionizing Transportation: Analyzing Solar Car Efficiency at EPIC Solar Farm 91
Afidatul Nadia Mok Hat, Ruzlaini Ghoni, Mohd Tarmizi Ibrahim, Ahmad Firdaus Zali, and Fuaad Mohamed Nawawi

Mathematical Modelling of Electromagnetic Transduction for Optimal Power Harvesting from Induction Motors 103
Ammar Husaini Hussian, Ruzlaini Ghoni, Mohd Tarmizi Ibrahim, Shaiful Rizalmee Wahid, Afidatul Nadia Mok Hat, Mohd Aizat Sulaiman, and Mohd Fadhil Ibrahim

Comparison of CNN and KNN Methods for Cataract Classification and Detection Based on Fundus Images 117
Farah Hanifah, Nur Alifia Azzahra, Yunendah Nur Fuadah, Alvian Pandapotan, Erni Yanthy, Rita Magdalena, and Sofia Saidah

Design of Photoplethysmography (PPG)-Based Respiratory Rate Measuring Device Through Peak Calculations 131
Ummul Muthmainnah, Willy Anugrah Cahyadi, Husneni Mukhtar, Muhammad Abdul Hakiim Al Fatih, and Denny Tri Sukmono

Automatic Obstructive Sleep Apnea Identification Using First Order Statistics Features of Electrocardiogram and Machine Learning 151
Aida Noor Indrawati, Nuryani Nuryani, Wiharto Wiharto, Diah Kurnia Mirawati, and Trio Pambudi Utomo

The Power Use of Power Spectrum Density for Measures of Cognitive Performance Based on Electroencephalography: Systematic Literature Review 167
Rahmaniyah Dwi Astuti, Bambang Suhardi, Pringgo Widyo Laksono, Novie Susanto, and Ainun Rahmansyah Gaffar

Enhancing Rice Leaf Disease Classification: A Combined Algorithm Approach for Improved Accuracy and Robustness 185
Apri Junaidi, Diao Qi, Chan Weng Howe, and Siti Zaiton Mohd Hashim

Improving Unbalanced Security X-Ray Image Classification Using VGG16 and AlexNet with Z-Score Normalization and Augmentation 205
Diao Qi, Apri Junaidi, Chan Weng Howe, and Azlan Mohd Zain

Mindfulness Intervention Affects Cognitive Abilities of Students: A Time–Frequency Analysis Using EEG 219
Trupti Taori, Shankar Gupta, Ramchandra Manthalkar, and Suhas Gajre

Wide Communication Coverage SpO_2 Monitoring Using Local Host HTML Web Page 235
I. Dewa Gede Hari Wisana, Nabila Surayya Saidah, Priyambada Cahya Nugraha, Moch Prastawa Assalim Tetra, Dessy Tri Wulandari, and Tetrik Fa'altin

Wide Communication Coverage ECG—Lead II Monitoring Using Local Host HTML Web Page 249
Priyambada Cahya Nugraha, Nurdiansyah Wahyu Bima Putra, I. Dewa Gede Hari Wisana, Moch Prastawa Assalim Tetra Putra, Riqqah Dewiningrum, and Divanda Natya Kirana

Effect of Higher Temperature at Metal Plate Based on Thickness and Hardness Material Using Ultrasonic Testing Method 263
Ahmad Anwar Zikri Othman, Kharudin Ali, Damhuji Rifai, Nazry Abdul Rahman, Zulfikri Salleh, Muhammad Ameen Wahab, Raja Siti Nur Adiimah Raja Aris, Johnny Koh Siaw Paw, Chong Tak Yaw, Jian Ding Tan, and Talal Yusaf

Enhancing Infant Safety: Performance Analysis of Deep Learning Method on Development Board for Real-Time Monitoring 285
Nugroho Budi Prasetyo, Dien Rahmawati, Wahmisari Priharti, and Muhammad Dhalhaz

Visualization of Data from a Multifunctional Surgical Device for Measuring Forces and Torques Using the Violin Diagram Method 303
Mikhail A. Solovyev, Andrey A. Vorotnikov, Andrey A. Grin, Yuri V. Poduraev, Anton Y. Kordonsky, and Oleg V. Levchenko

Analysis of Hyperparameters for Workout Movements Classification Using the Convolutional Neural Network Algorithm 317
M. Hasyim Abdillah Pronosumarto, Jiwa Sambhuwara, S. T. Koredianto Usman, and R. Yunendah Nur Fu'Adah

Developing a Patient Module-Based Mobile Application for Effective Self-isolation Management in COVID-19 Patients 335
Alfi Yusrotis Zakiyyah, Eka Miranda, Meyske Kumbangsila, Mediana Aryuni, Richard, and Albert Verasius Dian Sano

Improvements in the Imbalanced Hemogram Data Classification 347
Phuoc-Hai Huynh, Ngoc-Minh Nguyen, Trung-Nguyen Tran, and Thanh-Nghi Doan

Diagnose Skin Face Problems by Comparing Classification Algorithms 361
Marsya Ardini, Alzha Rizqie Kinanta, Vincensius Bunni Palagoro, Michael Alessandro Kevin Wibowo, and Aripin

Quickness Aspect Talent Identification (QATI) System Based on the Internet of Things 371
Rohmat Tulloh, Jordy Marchelino Lumban Gaol, Dery Rimasa, and Asep Mulyana

Sentiment Analysis on Service Quality of an Online Healthcare Mobile Platform Using VADER and Roberta Pretrained Model 383
Fairuz Iqbal Maulana, Puput Dani Prasetyo Adi, Dian Lestari, Agung Purnomo, and Daniel Anando Wangean

Internet of Medical Things: A Bibliometric Analysis of Research Publications from 2018–2022 395
Fairuz Iqbal Maulana, Dian Lestari, Puput Dani Prasetyo Adi, Mohammad Nazir Arifin, and Agung Purnomo

Integration of a Reverse Vending Machine Sensing System in Sorting and Detecting Plastic Bottle Waste 409
Juansah, Mohamad Ramdhani, and Dien Rahmawati

Corrosion Resistance Improvement of Mild Steel Using *Cocos Nucifera* Leaf Extract in Seawater 427
W. M. Wan Syahidah, R. Rosliza, and F. Atan

A Comprehensive Analysis of: A Systematic Review 437
Dian Lestari, Fairuz Iqbal Maulana, Agung Purnomo, and Puput Dani Prasetyo Adi

Enhancing Extractive Summarization in Student Assignments Using BERT and K-Means Clustering 453
Mamluatul Hani'ah, Vivi Nur Wijayaningrum, and Astrifidha Rahma Amalia

Topic Modeling of Raja Ampat Tourism on TripAdvisor Sites Using Latent Dirichlet Allocation 465
Dedy Sugiarto, Dimmas Mulya, Syandra Sari, Anung B. Ariwibowo, Is Mardianto, Muhammad Azka Aulia, Fitria Nabilah Putri, Ida Jubaidah, Arfa Maulana, and Alya Shafa Nadia

Bluetooth and Microcontroller Enabled Wireless Exposure Switch Development for X-ray Mobile Unit to Improve Radiation Protection 481
Akhmad Haris Sulistiyadi, Dwi Rochmayanti, and Ardi Soesilo Wibowo

State of the Art of Battery Swap Station Management System in Indonesia: An IoT-Based Prototype 493
Ridlho Khoirul Fachri, Rashad Abul Khayr, Muhammad Zakiyullah Romdlony, Aam Muharam, and Amin

Business Intelligence in Healthcare: A Review of Knowledge Structures and Level of Analysis 505
Agung Purnomo, Mega Firdaus, Fairuz Iqbal Maulana, Bigraf Triangga, Muchamad Indung Hikmawan, and Zahra Tazkia Nurul Hikmah

Assessing the Effectiveness of Mechanical Sensors for Respiratory Rate Detection 519
Dyah Titisari and M. Prastawa Assalim T. Putra

Obesity Prediction Approach Based Habit Parameter and Clinical Variable Using Self Organizing Map 531
Lilik Anifah, Haryanto, I. G. P Asto Buditjahjanto, R. R. Hapsari Peni Agustin Tjahyaningtijas, and Lusia Rakhmawati

Entropy-Based Analysis of Electromyography Signal Complexity During Flexion of the Flexor Carpi Radialis Muscle Under Varied Load Conditions 545
Katherine, Alfian Pramudita Putra, Angeline Shane Kurniawan, Dezy Zahrotul Istiqomah, Nisa'ul Sholihah, Khalid Ali Salem Al-Salehi, Khusnul Ain, Imam Sapuan, and Esti Andarini

Voice Features Examination for Parkinson's Disease Detection Utilizing Machine Learning Methods 559
Farika Tono Putri, Muhlasah Novitasari Mara, Rifky Ismail, Mochammad Ariyanto, Hartanto Prawibowo, Triwiyanto, Sari Luthfiyah, and Wahyu Caesarendra

Bone Drilling Vibration Signal Classification Using Convolutional Neural Network to Determine Bone Layers 577
Wahyu Caesarendra, Putri Wulandari, Kamil Gatnar, and Triwiyanto

Enhancing Accuracy in the Detection of Pneumonia in Adult Patients: An Approach by Using Convolutional Neural Networks 593
Lusiana, Arvita Agus Kurniasari, and Izza Fahma Kusumawati

Development of a Baby Incubator with a Vital Sign Monitoring Tool Based on Pan-Tompkins Method for Heart Rate Calculation 611
Bambang Guruh Irianto, Anita Miftahul Maghfiroh, M. Ikhsan Firmansyah, Abd. Kholiq, Syevana Dita Musvika, and Yuni Kusmiyati

Recent Advances and Challenges in 3D Printing of Prosthetic Hands ... 625
Triwiyanto, Sari Luthfiyah, Bedjo Utomo, I. Putu Alit Pawana, Wahyu Caesarendra, and Vijay Anant Athavale

A Review of 3D Printing Technology for the Development of Exoskeletons for Upper Limb Rehabilitation 643
Triwiyanto, Levana Forra Wakidi, Wahyu Caesarendra, Achmad Rizal, Abdussalam Ali Ahmed, and V. H. Abdullayev

A Review of Decomposition Methods for ECG-Derived Respiratory Signal Extraction: Principles, Performance, and Applications 665
Anita Miftahul Maghfiroh, Syevana Dita Musvika, Singgih Yudha Setiawan, Levana Forra Wakidi, and Farid Amrinsani

A Deep CNN-Based Approach for 10-Class with Two-Channel EMG Signal Classification 685
Triwiyanto, Endro Yulianto, Triana Rahmawati, and Rifai Chai

Multi Socket Transmission System Application with Advanced Encryption Standard Algorithm to Support Confidential Medical Data Security

Bita Parga Zen, Abdurahman, Anggi Zafia, Annisaa Utami, Iwan Nofi Yono Putro, and Faisal Dharma Aditama

Abstract Medical data information plays an important role in supporting the confidentiality of diagnosis data, treatment, patient health history, examination results, and laboratory examination results to assist doctors in making accurate diagnoses and choosing appropriate treatment; now, this data has been integrated into the internet. The information system in this application is needed to be more efficient and effective in supporting medical data security. To integrate a confidential medical data privacy system, the information system is equipped with an encryption security system on the NIK, Diagnosis, Amnesia, and Therapy tables to guarantee the privacy of the data sent. The prototype multi-socket network-based communication system, PHP with the Laravel Framework, uses the Advanced Encryption Standard (AES) encryption algorithm using the same Key and Initialization Vector (IV) for encryption and decryption. It was designed to support the medical data security system using the Laravel NET website programming language framework and MySQL database. The system has been built based on two access rights, doctor, and super admin, where doctors can enter patient diagnoses. AES encryption will be applied to all data when downloaded by a doctor. The super admin can input access rights as a doctor but cannot see the patient because it has been encrypted with the AES algorithm through medical records diagnosis. Because when the data is sent, this system is tested using black box testing with 34 black box tests to ensure the confidentiality of confidential patient data. The system with the AES 256 algorithm can be useful for patient data security systems in hospitals.

Keywords Encryption · Advanced encryption standards · Medical data · Websites

B. P. Zen (✉) · A. Zafia · A. Utami · F. D. Aditama
Institut Teknologi Telkom Purwokerto, Purwokerto, Indonesia
e-mail: bita@ittelkom-pwt.ac.id

Abdurahman
Universitas Sriwijaya, Palembang, Indonesia

I. N. Y. Putro
Badan Riset Dan Inovasi Nasional, Jakarta, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_1

1 Introduction

Various advances in information and communication technology science are developing so rapidly and broadly that this provides great support for progress, security, and resilience for the protection of a country [1]. The development of technology and information also creates various network-based threats, utilizes information, and can carry out digital and cyber wars [1]. As reported in the media, Bisnis.com. Millions of hospital patient data were allegedly leaked and sold on dark forums. The hacker claims the data comes from "the centralized server of the Indonesian Ministry of Health." Uploaders on the forum attach a sample of 6 million data containing, among other things, the patient's full name, hospital, patient photo, COVID-19 test results, and X-ray scan results [2]. This, of course, can affect the occurrence of cracks in a medical data information system that plays an important role in supporting the confidentiality of data on diagnosis, treatment, patient medical history, examination results, and laboratory test results can help doctors make an accurate diagnosis and choose the right treatment [3].

Currently, these data have been integrated into the internet, with advances in various information technology knowledge, and cyber security systems are currently developing so rapidly and provide great support for progress, data security, with technological advances in the field of cyber [4], especially on websites and databases can become a threats in theft of medical data and information so that it is necessary to have a medical data security system to deal with data theft [5]. A medical data security information system application is needed so that the system can run very well from various types of tests that are carried out well [6] and provide great depth and a wide selection of tools for use in website applications in the security testing process by covering the stages of focusing on the process used in designing, improving, and managing website security [7].

Cryptography is a technique for encoding information, so unauthorized parties cannot read or see it. There are three types of cryptography: symmetric cryptography, asymmetric cryptography, and hash functions [8]. Symmetric cryptography is an encryption algorithm that utilizes the same key for each encryption and decryption process, while asymmetric cryptography utilizes different keys for the encryption and decryption process [9]. The hash function is a one-way algorithm that does not perform the decryption process. The symmetric algorithm has a higher speed than the asymmetric algorithm, so the symmetric algorithm is suitable for encrypting large amounts of data, besides that the symmetric algorithm is also strong because of complex arithmetic calculations [10]. Based on the need for confidentiality of medical data, this scientific research formulated an interoperability solution by designing a website system based on the Laravel framework using cryptographic algorithms as medical data security so that data integrity is guaranteed [11].

This study used the Advanced Encryption Standard (AES) encryption algorithm for data transmission security. The AES encryption algorithm is a symmetric encryption with the same public and private keys [12]. The AES encryption algorithm was chosen in this study because the AES encryption algorithm is considered to have

high speed and strength in the encryption and decryption process [13]. In addition, the AES algorithm is a symmetrical algorithm that is quite safe now compared to other algorithms [14]. These factors are balanced with information needs that can exchange information efficiently and effectively because in exchanging information, speed is a crucial factor [15].

The system uses the Laravel framework with text encryption via OpenSSL using AES-256 with a key configuration that will contain information about sending data about patients, communication data with practicing doctors, communication of medical services, communication about medical conditions and symptoms, information about medicines, and information about medical procedures:

List of available medical procedures, procedure description, procedure time, and procedure fees, contact information: clinic or hospital web address, phone number, and email, important to protect medical data and ensure only authorized persons have access to this information because all data transmissions are properly encrypted, with the hope that this system can maintain the confidentiality of data from illegal theft because patient data is a vital thing that must be kept safe.

2 Materials and Methods

The tools and materials used in this study are listed in Table 1. In this study, the analysis phase was carried out on encrypted and decrypted data by comparing the results and whether the initial plaintext with the final plaintext was appropriate. By implementing the AES encryption algorithm, text data in ciphertext (data after being encrypted) has a bit length different than plaintext (data before being encrypted) [16]. This affects the time needed to transmit data. The different algorithms will have different results, such as packet size, contents of encrypted data, etc. Data in the form of cipher text contains information, namely the destination IP address and the port used.

The algorithm used in this study is the Advanced Encryption Standard 256-bit key encryption algorithm which uses the number of keys (Nk) = 8, block size (Nb) = 4, and performs 14 rounds (Nr). The 256-bit AES algorithm was chosen in this study because the key has a longer mathematical calculation by making 14 rounds of transformation. The 256-bit key processes the most transformations than the 128-bit and 196-bit keys [17] so it is calculated to be stronger than the AES algorithm

Table 1 Software requirements

Name	Information	Qty	Function
Framework Laravel	Programming language	1	Website development
Visual studio code	Application coding	1	*Web*
MySQL	Database	1	Database
Chrome	Search engine	1	Interface

that uses other keys. The AES algorithm utilizes the same key for the encryption and decryption process and consists of four types of transformation bytes: AddRoundKey, SubBytes, ShiftRows, and MixColumns [18].

The encryption and decryption process uses the key and Initialization Vector (IV) agreed upon at the beginning. An initialization vector is a number used with the key to make it more difficult for crackers to break the encryption pattern when using a dictionary attack, In Fig. 1 explains the AES Encryption algorithm workflow in the development process of this website, first the doctor inputs the login and input the password, if the user enters patient data, disease, history and so on, every data input will be encrypted by the AES 256 algorithm, for this input the admin can only see the patient and select the patient, for complete disease history data the admin cannot see, and only the doctor can see the patient's disease, and the admin must ask for the doctor's approval to see the patient's personal data.

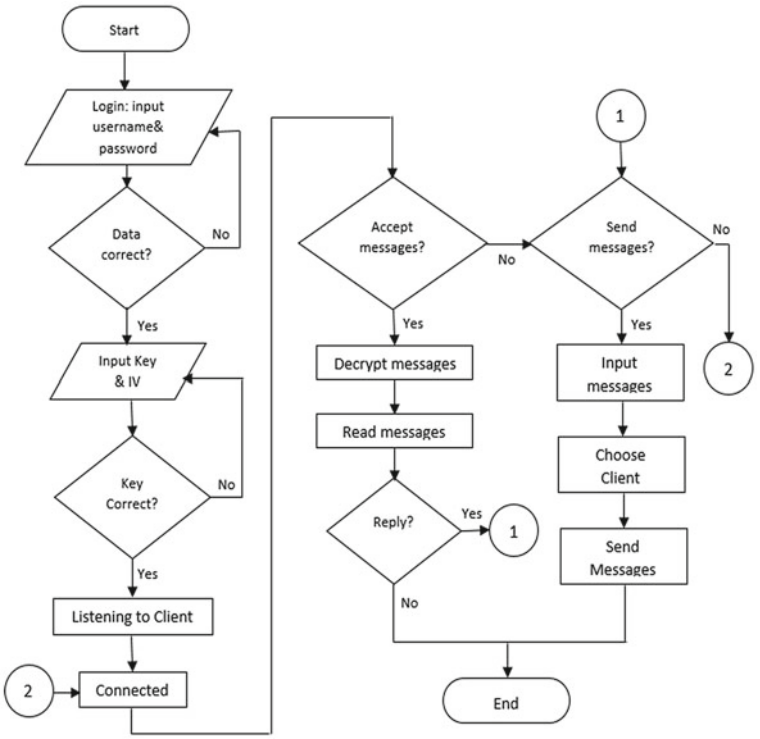


Fig. 1 Server program flow

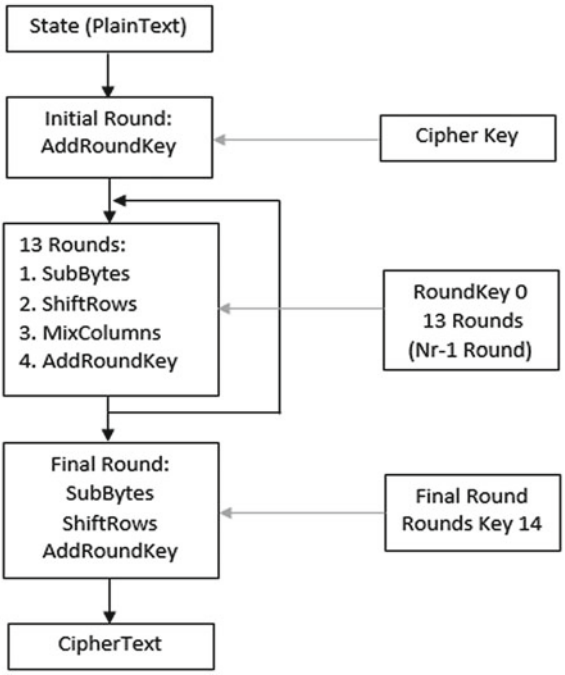


Fig. 2 AES encryption process

2.1 Encryption Process

In Fig. 2, explains the flow the encryption process is done by inputting the cipher key and then transforming it. The first step is initializing the AddRoundKey round by inputting the agreed 256-bit key. After that, 13 rounds of SubBytes transformation, ShiftRows transformation, MixColumns transformation, and AddroundKey transformation are performed and repeated up to 13 rounds. Then the last or 14th SubBytes, ShiftRows, and AddroundKey transformations are performed before the ciphertext output is output. For Cipher and Inverse Cipher, the AES algorithm uses a round function that consists of four different byte-oriented transformations: (1) substitution of each byte using a substitution table or S-box, (2) shifting state array rows with different offsets, (3) mixing the data in each column of the state array, and (4) adds a Round Key to the process state [19].

2.2 Descryption Process

Figure 3 explains the decryption process workflow, The decryption process transforms InvShiftRows, InvSubBytes, InvMixColumn, and AddRoundKey. The

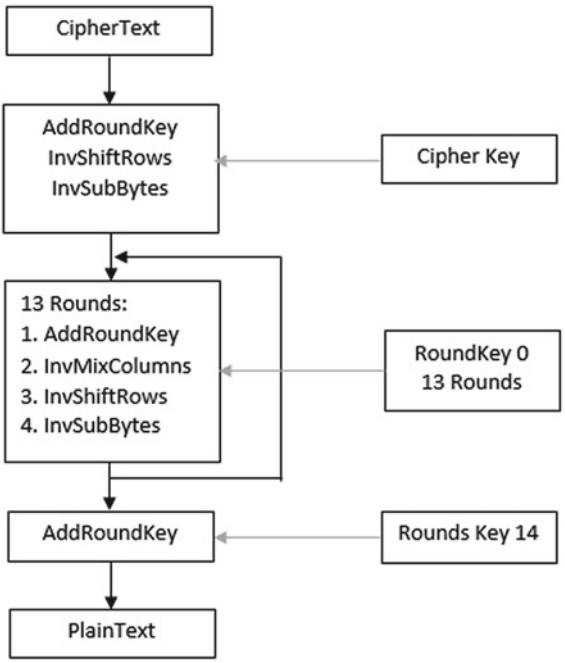


Fig. 3 AES decryption process

InvShiftRows transformation has a reverse process with the ShiftRows transformation.

The ShiftRows transformation shifts bits to the left, while the InvShiftRows transformation shifts bits to the right. The InvSubBytes transformation is a transformation that is the opposite of the SubBytes transformation. In this transformation, each component in the state is described using the Inverse S-Box table; InvMixColumns is the opposite of the MixColumn() transform. Each column in the state is multiplied by the multiplication matrix of the AES encryption system. Figure 3 also explains the script the process of implementing decryption.

2.3 Blackbox Testing

In this study, black box testing was conducted to test the functionality of the system and ensure that the function can run according to the design of the system. It is done by performing procedures or testing steps to determine whether the test results are in line with expectations or not [20]. In this study, black box testing was conducted to test the functionality of the system and ensure that the function can run according to the design of the system. It is done by performing procedures or testing steps to determine whether the test results are in line with expectations or not.

3 Result

In result the system view below is a display of medical data as super admin using the Laravel framework, where this data is encrypted with the Advanced Encryption Standard (AES) algorithm; in the image below, the access rights as super admin have access to print patient data which only has access Name, No BPJS and NIK, but cannot enter patient data for diagnostic data, amnesia data and therapy as superadmin does not have access to view patient diagnosis data, because this data has been encrypted properly using the Advanced Encryption Standard algorithm with a 256-bit key encryption algorithm which uses the number of keys (Nk) = 8, block size (Nb) = 4, and performs 14 rounds (Nr) so superadmin can't see patient history because it's encrypted properly. Where in the image displayed below, there is a dashboard, patient, and diagnosis, on the dashboard, displays total doctor records and total patient records. Figure 4 displays patient data in which there are entries for Name, NIK, Place, and Date of Birth, NIK, HP No, BPJS, Name of Head of Family, Occupation, Gender, Religion, Address, History of Allergies, History of Trauma, History of Disease, History of Hereditary Diseases This data can be printed by doctors, patient data; doctors can also print patient data.

In Fig. 5, it can be seen that the superadmin cannot see the patient's diagnosis, amnesia, or tension because everything has been properly encrypted using AES 256; the superadmin can only see the patient's biodata.

In Fig. 6, at diagnosis only the doctor's account can see a summary of diagnoses, amnesia, therapy, allergies and a summary of the patient's illnesses that become doctor's records, besides that doctors can also create new patient data as shown in the image.

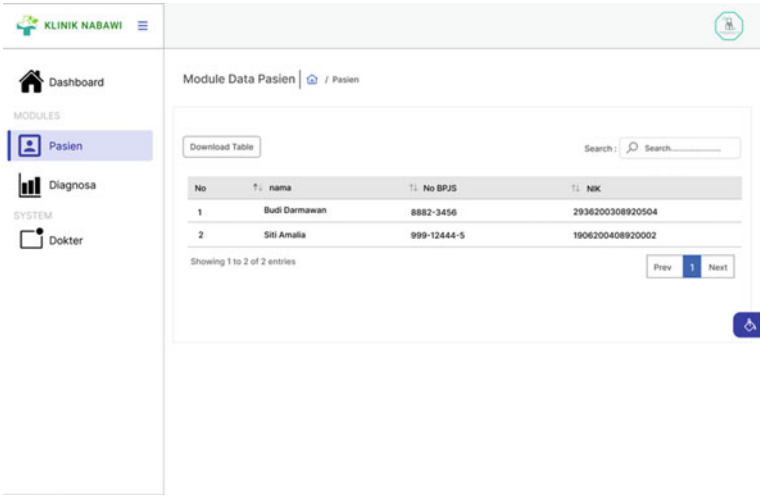


Fig. 4 Display of medical data as super admin

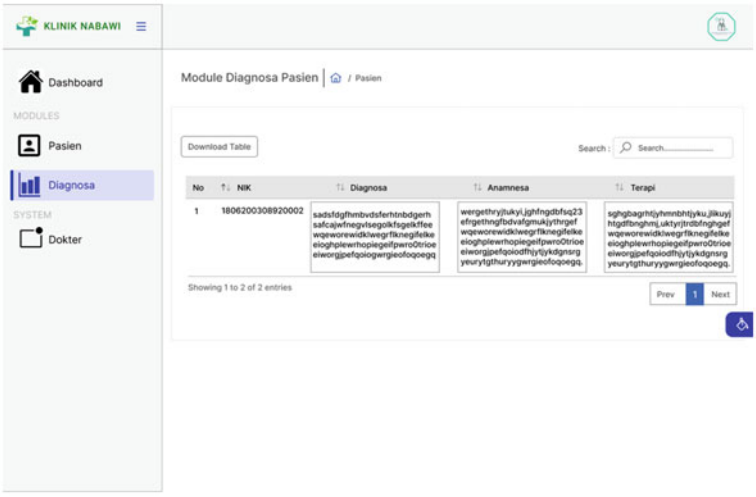


Fig. 5 Patient diagnostic confidential data encryption results

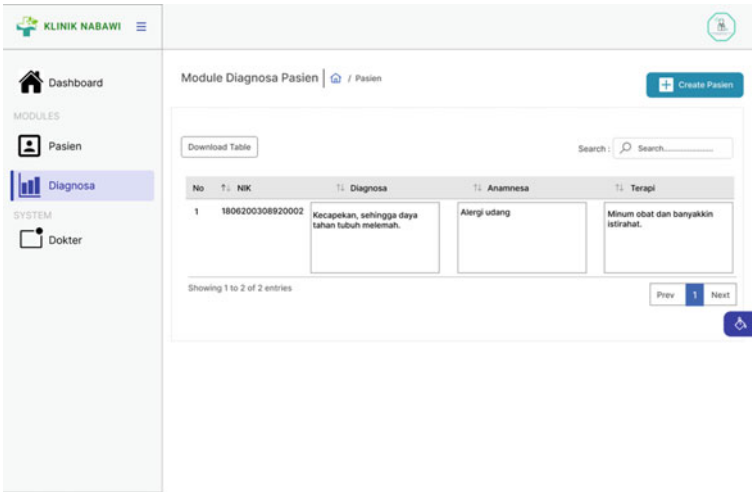


Fig. 6 Patient diagnostic module

This system uses black box testing which can be seen in Table 2 This system consists of 2 users, namely doctors s and super admins. Testing is carried out on the login menu, menu on superadmin and doctor menu. From these tests, the system can run well according to the flow of given scenarios from the input and output processes.

Blackbox Testing. Attached below are the ID codes that are used as a function of each action in the system. For ID L, it is used to login to Super Admin and Doctor when entering the system; for ID SA, it is used as a symbol for access as super admin;

Table 2 Black box testing

ID	Test step	Expected results	Results obtained
L001	1. Opens the login page 2. Enter *username* "doctor" 3. Enter *password* 4. Press login	The system displays the doctor's *user page*	The system displays the website system doctor *user page*
L002	1. Opens the login page 2. *The username* and *password* form is not filled in 3. Press the login button	The system displays an alert notification "*username not filled in*" and "*Password not entered*"	The system displays an alert notification "*username not filled in*" and "*password not filled in*" when clicking the login button
D001	1. Open patient page 2. Displays patient data table 3. Press the *Create patient data button* to add patient data	The system displays the form added patient data	The system displays the patient data form when pressing the *create* patient data button
D002	1. Open patient page 2. Enter patient data 3. Press the save button	The system stores data and displays data in the patient table, if the save button is pressed	The system displays patient data that has been added to the patient table
D003	1. Open patient page 2. Hit the edit button	The system will open an edit form according to the selected data	The system displays the patient edit form
SA01	1. Open the doctor's page 2. Displays the doctor's data table 3. Press the *Create doctor data button* to add doctor data	The system displays the doctor data added form	The system displays the doctor's data form when pressing the *create* data doctor button
SA02	1. Open the doctor's page 2. Enter doctor data 3. Press the save button	The system saves data and displays data in the doctor's table, if the save button is pressed	The system displays doctor data that has been added to the doctor table
SA03	1. Open the doctor's page 2. Hit the edit button	The system will open an edit form according to the selected data	The system displays the doctor's edit form

(continued)

Table 2 (continued)

ID	Test step	Expected results	Results obtained
DSA01	1. Opens the diagnostic by superadmin page 2. Displays an encrypted diagnostic data table	The system displays a patient diagnosis data table	The system displays encrypted patient diagnostic data
DSA02	1. Opens the diagnostic by superadmin page 2. Displays the patient diagnostic data download button	If the download button is pressed for patient diagnostic data, the data will be downloaded	The system has successfully downloaded the patient's diagnostic data properly
DT01	1. Opens the doctor's diagnosis page 2. Displays the diagnostic data table	The system will display a patient diagnosis data table	The system can display a patient diagnosis data table
DT02	1. Opens the doctor's diagnosis page 2. Displays the diagnosed data table 3. Hit the edit button	The system will open an edit form according to the selected data	The system displays the patient diagnosis edit form and presses the save button to save the data
DSAP1	1. Open patient page by superdmin or doctor 2. Displays the patient data download button	If the patient data download button is pressed, the data is downloaded	The system successfully downloads patient data properly

for ID D, it is performed by a doctor; for the DSA, ID is used as a diagnostic menu by super admin, for the DT ID as a diagnostic menu by a doctor, here is the complete information below.

4 Discussion

The implementation result of this system is using black box testing tables which can be seen in Table 2. This system consists of 2 users, namely doctors as experts and admins. Then, test on the login menu which can be seen on IDs L001 and L002. Where in this test produces every input and output process, and the features owned can run

well according to the flow of the given scenario. Black box testing from the doctor's side can be seen in ID D001, D002, D003, DT01 and DT02 where in this test produces every input and output process, and the features owned can run well according to the flow of the given scenario. Black box testing from the super admin side can be seen in SA01, SA02, SA03, DSA01, DSA02 and DSAP1 where in this test produces every input and output process, and the features owned can run well according to the flow of the given scenario. Some studies that use encryption such as those conducted by [21]. This research uses ZeroMQ, then ZeroMQ sends messages openly without securing the messages sent. This is evidenced by research that states that ZeroMQ does not have a layer of security for sending messages over the network and direct observation on ZeroMQ message packets using the Wireshark application. In this study using methods of securing and authenticating message delivery using AES encryption (Advanced Encryption Standard) Mode CBC (Cipher Block Chaining) combined with simple authentication methods. Research conducted by [22] performs the implementation of the AES algorithm for all file types. Research conducted by [23] using AES Algorithm on internet of things networks to support smart healthcare.

This research designed a prototype multi-socket network-based communication system using Advanced Ency Standard (AES) encryption algorithms to support medical data security systems designed using PHP programming language .NET framework and SQL Server databases. The system is built based on a client/server design where there is one server and several clients. AES encryption algorithm is used in multi-socket network transmission systems as a security system to ensure data confidentiality when data is transmitted. AES encryption algorithm is considered to have high speed and strength compared to other encryption algorithms. The AES encryption algorithm uses the same Key and Initialization Vector (IV) for the encryption and decryption processes. This research is a new innovation that discusses socket programming by utilizing the AES encryption algorithm as a security system. Research can use a variety of methods, for example case-based reasoning methods [24, 25].

5 Conclusion

This medical data application has been encrypted with the AES 256 algorithm. This answers the background problem and is a reference for policymakers regarding the importance of medical data security from cyber-attacks. This is very personal and sensitive information on a system that has been created using a combination of AES encryption with libraries. This study used the Advanced Encryption Standard (AES) encryption algorithm for data transmission security. The AES encryption algorithm is a symmetric encryption with the same public and private keys. The AES encryption algorithm was chosen in this study because the AES encryption algorithm is considered to have high speed and strength in the encryption and decryption process. Future research could be developed using it to diagnose diseases, rather than just the process of encryption and decryption.

References

1. Aceto G, Persico V, Pescapé A (2018) The role of information and communication technologies in healthcare: taxonomies, perspectives, and challenges. J Netw Comput Appl (Academic Press) 107:125–154. https://doi.org/10.1016/j.jnca.2018.02.008
2. Newswire, "Jutaan Data Pasien RS Indonesia Bocor, Ini Kronologi Lengkapnya," https://teknologi.bisnis.com/read/20220107/84/1486327/jutaan-data-pasien-rs-indonesia-bocor-ini-kronologi-lengkapnya, Jakarta, January 07, 2022
3. Iyengar K, Upadhyaya GK, Vaishya R, Jain V (2020) COVID-19 and applications of smartphone technology in the current pandemic. Diabetes Metab Syndr Clin Res Rev 14(5):733–737. https://doi.org/10.1016/j.dsx.2020.05.033
4. Li Y, Liu Q (2021) A comprehensive review study of cyber-attacks and cyber security; emerging trends and recent developments. Energy Rep 7:8176–8186. https://doi.org/10.1016/j.egyr.2021.08.126
5. Javaid M, Haleem A, Singh RP, Suman R (2023) Towards insighting cybersecurity for healthcare domains: a comprehensive review of recent practices and trends. Cyber Secur Appl 1:100016. https://doi.org/10.1016/j.csa.2023.100016
6. Munoz-Gama J et al (2022) Process mining for healthcare: characteristics and challenges. J Biomed Inform 127. https://doi.org/10.1016/j.jbi.2022.103994
7. Zen BP, Gultom RAG, Reksoprodjo AHS (2020) Analisis security assessment Menggunakan metode penetration testing Dalam Menjaga Kapabilitas Keamanan Teknologi Informasi Pertahanan Negara. J Teknol Penginderaan 2(1):105–122
8. Liestyowati D (2020) Public key cryptography. J Phys Conf Ser 1477(5). https://doi.org/10.1088/1742-6596/1477/5/052062
9. William P, Choubey, A, Chhabra GS, Bhattacharya R, Vengatesan K, Choubey S (2022) Assessment of hybrid cryptographic algorithm for secure sharing of textual and pictorial content
10. Zhang Q (2021) An overview and analysis of hybrid encryption: the combination of symmetric encryption and asymmetric encryption. In: 2021 2nd international conference on computing and data science (CDS), pp 616–622
11. Akbulut S et al (2023) Designing a private and secure personal health records access management system: a solution based on IOTA distributed ledger technology. Sensors 23(11). https://doi.org/10.3390/s23115174
12. Parenreng JM, Mustari SM, Wahid A (2022) E-mail security system using El-Gamal hybrid algorithm and AES (advanced encryption standard) algorithm. Internet Things Artif Intell J 2(1):1–9. https://doi.org/10.31763/iota.v2i1.510
13. Nida HS, Bura RO, Abdurahman (2021) Multi-network transmission using socket programming to support command and control systems
14. Al-Shabi MA (2019) A survey on symmetric and asymmetric cryptography algorithms in information security. Int J Sci Res Publ 9(3):8779. https://doi.org/10.29322/ijsrp.9.03.2019.p8779
15. Naser Alsuqaih H, Hamdan W, Elmessiry H, Abulkasim H (2023) An efficient privacy-preserving control mechanism based on blockchain for E-health applications. Alexandria Eng J 73:159–172. https://doi.org/10.1016/j.aej.2023.04.037
16. Anwar N, Abduh M, Santosa NB (2018) Komparatif performance model Keamanan Menggunakan Metode Algoritma AES 256 bit dan RSA. 2(3):783–791. http://jurnal.iaii.or.id
17. Saha R, Geetha G, Kumar G, Kim TH (2018) RK-AES: an improved version of AES using a new key generation process with random keys. Secur Commun Netw 2018. https://doi.org/10.1155/2018/9802475
18. Muhammad Abdullah A, Muhamad Abdullah A (2017) Advanced encryption standard (AES) algorithm to encrypt and decrypt data application of petri nets in computer networks view project call for papers view project advanced encryption standard (AES) algorithm to encrypt and decrypt data. https://www.researchgate.net/publication/317615794
19. Manjula G, Mohan HS (2016) Constructing key dependent dynamic S-Box for AES block cipher system

20. Sasmito GW, Nishom M (2020) Testing the population administration website application using the black box testing boundary value analysis method. In: 2020 IEEE conference on open systems, ICOS 2020, pp 48–52. https://doi.org/10.1109/ICOS50156.2020.9293645
21. Made Sukarsa I, Made I, Pradana R, Wira Buana P (2017) Implementasi Enkripsi dan Otentikasi Transmisi Data ZeroMQ Menggunakan advanced encryption standard. RESTI 1(3):1149–1156
22. Yuniati V, Indriyanta G, Rachmat CA (2011) Enkripsi Dan Dekripsi Dengan Algoritma Aes 256 Untuk Semua Jenis File. J Inform 5(1). https://doi.org/10.21460/inf.2009.51.69
23. Rachmayanti A, Wirawan W (2022) Implementasi Algoritma advanced encryption standard (AES) pada Jaringan internet of things (IoT) untuk Mendukung smart healthcare. J Tek ITS 11(3). https://doi.org/10.12962/j23373539.v11i3.97042
24. Utami A, Suyanto Y, Sihabuddin A (2021) Reccomendations on selecting the topic of student thesis concentration using case based reasoning. IJCCS (Indonesian) J Comput Cybern Syst 15(1):1. https://doi.org/10.22146/ijccs.58919
25. Wahyudi E, Hartati S (2017) Case-based reasoning untuk Diagnosis Penyakit Jantung. IJCCS (Indonesian) J Comput Cybern Syst 11(1):1. https://doi.org/10.22146/ijccs.15523

The Application of Virtual Reality Using Kinect Sensor in Biomedical and Healthcare Environment: A Review

Henry Candra, Umi Yuniati, and Rifai Chai

Abstract Virtual Reality (VR) and the Kinect sensor, as promising tools in biomedical research, offers diverse applications in medical education, rehabilitation, diagnostics, and health research. The problem statement highlights the demand for innovative solutions and introduces VR and Kinect as potential transformative technologies. This review analyzes the importance of these technologies, their contributions, and future potential. It stands out by evaluating various Kinect-based systems in medical settings. By highlighting distinct features, advancements, and limitations, it provides guidance for future research. Relevant literature was gathered from databases such as Google Scholar, IEEE Xplore, and PubMed. The results showcase a wide range of applications, including patient autonomy, stroke rehabilitation, diagnostics, and monitoring. Despite challenges in accurate movement tracking, integration into clinical settings, and limited generalizability of findings due to small sample sizes, VR and Kinect show potential for revolutionizing healthcare delivery and improving patient outcomes. Their adaptability, affordability, and immersive nature of these technologies offer promising avenues for personalized interventions, remote healthcare, training, and enhanced patient engagement. As these technologies evolve, continued research and development are crucial to optimize their impact in shaping the future of healthcare.

Keywords Virtual reality · Kinect sensor · Biomedical fields · Healthcare solutions

H. Candra · U. Yuniati (✉)
Electrical Engineering Department, Faculty of Industrial Technology, Universitas Trisakti, Jakarta, Indonesia
e-mail: 162012200005@std.trisakti.ac.id

R. Chai
School of Science, Computing and Engineering Technologies, Swinburne University of Technology, Melbourne, Australia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_2

1 Introduction

Virtual reality (VR) has become a promising biomedical research tool, showing great potential in recent years. Its applications span various areas within the field, such as in medical educational, surgical simulation, and therapy [1]. For instance, VR instruments like headsets and motion controls have been employed in rehabilitation therapy to aid patients recovering from stroke [2] and brain injuries. By allowing individuals to practice motor skills in a secure virtual environment, these systems contribute to accelerating the recovery process and improve motor function [3].

In addition to rehabilitation, VR finds utility in medical education. Medical and healthcare students can utilize VR headsets and other resources to simulate medical procedures and explore human anatomy virtually, such as surgical simulation [4]. This immersive approach enhances their training experience, providing a more realistic environment to develop their medical skills [5]. Furthermore, VR is valuable in medical diagnostics, enabling doctors to make more precise diagnoses by using data from medical imaging [3]. The mechanisms allow doctors to identify health problems and develop more effective treatment plans quickly.

Moreover, VR is vital in health research, particularly in psychology, neurology, and psychiatry. Researchers can create controlled environments using VR headsets and observe participants' responses and behaviors in simulated situations. These experiments facilitate the exploration and development of novel interventions for various health problems.

The Kinect sensor technology, as a virtual reality device, stands out for its affordability, user-friendliness, and potential for advancement in medical and healthcare applications [6] in recent years. The advent of the Kinect and associated libraries has led to increased research interest on practical applications [7].

This review aims to comprehensively analyze the role of these technologies, the contributions they make to healthcare, and their implications for the future. This study distinguishes itself by conducting a systematic comparison and analysis of a diverse range of Kinect-based systems applied to biomedical and healthcare environment. By highlighting the unique features, innovations, and limitations of each approach, this review provides a comprehensive overview that can guide future research and development efforts.

2 Virtual Reality (VR) and Kinect Sensor

2.1 Definition and History of VR

Virtual reality is a computer program that enables individuals to connect and engage with computer-generated surroundings, replicating real-life experiences and involving all the senses. It offers interactive and immersive experiences, commonly known as the two I's. However, there exists a lesser known third aspect of virtual

reality. Virtual reality possesses applications that provide practical solutions to real-world challenges beyond being a medium or advanced user interface. Virtual reality professionals develop these applications. The effectiveness of a particular application, or how well a simulation performs, heavily relies on the human imagination—the third "I" of virtual reality. Thus, virtual reality represents immersion, interaction, and imagination. Vision in virtual reality also pertains to the mind's ability to perceive things that do not exist [8].

Virtual reality has a history spanning over 40 years. In 1962, Morton Heilig patented the Sensorama Simulator. This first virtual reality video arcade offered a multi-sensory experience with 3D video, motion, colour, sound, aromas, wind effects, and a vibrating seat. Sutherland's work in 1966 introduced the concept of head-mounted displays (HMDs) and the use of computer-generated scenes instead of analog images. Brooks and his colleagues later simulated force fields and collision forces using robotic arms, paving the way for today's haptic technology. The military's interest in digital simulators drove classified research, and NASA developed the Virtual Visual Environment Display (VIVED) in 1981, influencing modern VR headsets. NASA's integration of computers and trackers created the first VR system, while Scott Fisher's contributions enhanced interaction with sensing gloves and virtual sound sources. These advancements spurred international conferences and firmly established virtual reality in the scientific and engineering communities.

VPL Inc., led by Jaron Lanier, emerged as the first to sell VR products. They introduced the groundbreaking DataGlove, which revolutionized computer interaction with its fiber-optic sensors enabling gesture-based input in 1987. However, the DataGlove was expensive and lacked tactile feedback. In 1989, Nintendo released the more affordable PowerGlove, although its downfall was the limited game support. VPL also introduced the EyePhones, commercial head-mounted displays, in the late 1980s, but they suffered from low resolution and high prices. Integration challenges persisted, prompting VPL to develop the turnkey VR system, RB2 Model 2. Division Ltd. introduced the integrated VR workstation, Vision, and the consequential Provision 100 in 1992. Sense8 Co.'s WorldToolKit and Dimension International's Virtual Reality Toolkit (VRT3) addressed the software development challenges in the 1990s.

Early virtual reality (VR) hardware faced several challenges in the 1990s. The VR market was small, with most pioneering companies lacking resources for product improvements and relying on private capital. Unrealistic public expectations and funding limitations led to the disappearance of many VR companies. However, a small group of scientists continued VR research. Significant advancements in PC hardware, such as faster CPUs and graphics accelerators, facilitated VR's rebirth in the late 1990s. The performance of PC graphics matched or exceeded high-end SGI graphics supercomputers by 2001, thanks to rapid technological advances. Additionally, breakthroughs in VR interfaces, including lightweight and higher-resolution LCD-based HMDs and the introduction of large-volume displays, further contributed to the growth of the VR market.

2.2 VR Technology and Its Components

Virtual reality (VR) systems have five main components: VR Engine, Input/Output Devices, Software and Database, User, and Task [8]. The block diagram in Fig. 1 illustrated the architecture of VR [9]. VR Engine is the core software framework that powers VR experiences by rendering virtual environments, managing physics simulations, integrating tracking data, and enabling real-time interactivity. Input devices capture user interactions and movements, while output devices provide sensory feedback. Examples include motion controllers, data gloves, positional trackers, head-mounted displays (HMDs), audio systems, and haptic feedback devices. Software and databases are essential for developing, customizing, and deploying VR applications. The developments include content creation tools, simulation software, user interface systems, and APIs. A database or content management system stores virtual assets like 3D models, textures, sounds, and user data. The user is the participant who engages with the VR system through input and output devices, experiencing a sense of presence and interaction. User comfort, safety, and overall experience are crucial considerations. The task is the purpose and objectives of using VR, defining specific activities, simulations, or experiences the system aims to provide. VR applications span various domains, such as gaming, entertainment, education, training, and scientific research. The task component guides the design and development of the virtual environment and user interactions to support the intended use case.

Through the seamless integration of these five components, virtual reality systems generate attractive and interactive encounters, transporting users to simulated realms where they can explore, interact, and engage in manners surpassing the physical world's limitations.

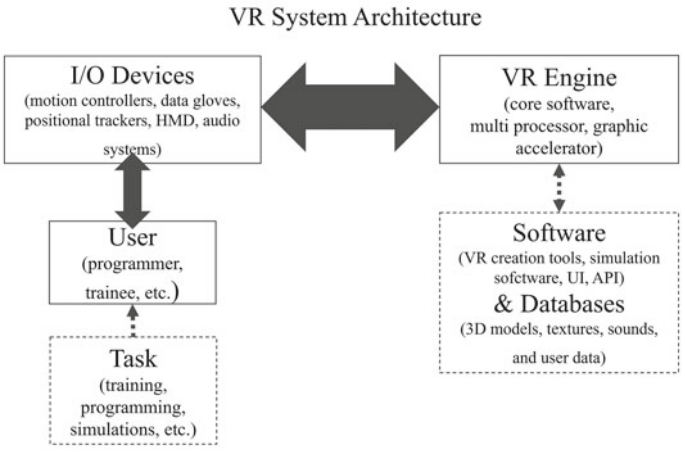


Fig. 1 The five components of a VR system

2.3 The Role of Kinect Sensor in VR

The Kinect sensor is a motion-sensing input device initially developed by Microsoft for gaming on the Xbox console. Kinect sensor V1, released in 2010, utilized infrared depth sensing, an RGB camera, and microphones to capture motion and audio data [10]. It allowed users to control games and menus through body movements and voice commands. Kinect Sensor V2, introduced in 2014, offers improved depth and image sensors, which may enhance its accuracy in assessing postural control and balance impairments. Both versions employed skeletal tracking technology, enabling immersive interactions through gestures, voice, and facial recognition. However, Kinect V2 provided better performance and advanced features, making it the preferred choice for users and developers in virtual reality, gaming, and other applications [11].

While it was primarily designed for gaming, the Kinect sensor has also found applications in VR systems. Known as a depth camera, it revolutionizes robot perception worldwide, replacing traditional cameras and range finders. Its depth-sensing capabilities enable detailed environment understanding and map generation, setting it apart from other vision systems. The Kinect sensor's ability to capture high-quality images and depth information about the environment at an affordable price has led to its widespread adoption in numerous research projects [12].

Research by El-laithy et al. aims to integrate the Kinect sensor into an autonomous ground vehicle, the Unmanned Utility Robotic Ground Vehicle (UURGV), to enhance its capabilities. The UURGV can autonomously cut grass and navigate outdoor areas using a laser scanner sensor and GPS. However, GPS navigation is ineffective indoors or near solid concrete structures. They propose integrating the Kinect sensor with other sensors, such as an inertial measurement unit (IMU), to optimize indoor navigation. By utilizing the Kinect's depth camera, the robot can detect approximate distances to objects and navigate around them while also gaining 3D perception [12].

Correa et al. project introduces a perception system for autonomous navigation of mobile surveillance robots in indoor environments. The method comprises two components: a reactive navigation system that uses the Kinect distance sensor to avoid obstacles and an artificial neural network (ANN) trained on Kinect data to identify various environment configurations. By combining reactive and deliberative behaviours, the robot can navigate using a topological map represented as a graph. The system shows promising outcomes regarding autonomous mobile robot navigation, including the ability to operate in dark environments, and its effectiveness was verified using a Pioneer P3-AT robot [13].

The paper by Eric and Jang focuses on using the Kinect depth sensor, a low-cost range sensor, to detect objects and measure their distance, providing an affordable and reliable option for computer vision. The results demonstrate that the Kinect sensor and segmentation techniques can effectively detect objects and accurately measure their distance. This research makes it suitable for obstacle avoidance and other applications that require a vision sensor, not only for unmanned vehicles but

also for manned vehicles to alert drivers. Computer vision plays a crucial role in vehicles without human intervention [14].

A common use of the Kinect sensor in VR is full-body tracking. It allows users to see their body movements replicated in the virtual world, enhancing the sense of presence and embodiment. The method can be precious in VR applications focused on physical training, rehabilitation, or social interactions. Additionally, the Kinect sensor's microphone array enables voice recognition and voice commands, allowing users to interact with the VR environment using speech. The system can enhance the usability and hands-free nature of VR experiences.

Clark et al. conducted a study to evaluate and compare the accuracy and consistency of kinematic data collected from a marker-based 3D motion analysis system and the Kinect Sensor V2 across different static and dynamic balance evaluations. The results showed excellent concurrent validity for trunk angle data in active tasks and anterior–posterior range and path length in static balance tasks, indicating the potential of the Kinect Sensor V2 as a reliable and valid tool for assessing specific aspects of balance performance [15].

3 Applications of VR and Kinect Sensor in Biomedical and Health Fields

Virtual reality (VR) has shown significant potential in the biomedical field, offering a range of applications. In these domains, using virtual reality (VR) involves the implementation of computer-generated environments and simulations that deeply engage individuals, aiming to improve medical research, education, diagnostics, healthcare services, treatment, and rehabilitation. By leveraging VR technology, biomedical professionals can explore and interact with virtual representations of anatomical structures, physiological processes, and medical scenarios.

A systematic review methodology was employed to gather relevant literature and projects related to the applications of VR and Kinect sensors in healthcare. Databases such as Google Scholar, IEEE Xplore, and PubMed were searched for articles and studies published between 2013 and 2022. Keywords included "Virtual Reality," "Kinect Sensor," "Healthcare," "Biomedical," "Medical Applications," and related terms. Studies were selected based on relevance to the topic and their contributions to the field.

For example, the paper by Kassem et al. introduces a smart digital medical bed called MedBed, designed to address critical issues in hospitals related to timely nurse intervention and patient independence. MedBed is intended to give patients autonomy and enable them to take necessary actions when nurses are unavailable. The bed incorporates features ranging from basic to highly impactful on the patient's vital status. It allows voice commands from the patient and communicates as part of the Internet of Things (IoT) through a user-friendly smartphone application. A

database is included to track activities such as medication and reminders. MedBed can be implemented in medical centers with minimal infrastructure changes [16].

Hesham A. Alabbasi et al. in their research [17], aim to utilize facial emotion recognition and brain activity to detect facial expressions and identify the corresponding brain regions in stroke patients with limited verbal communication. The proposed method involves tracking the patient's face and using the recorded coordinates to classify facial expressions through a neural network in Matlab. The results are displayed on the screen, and while initially focused on stroke patients, the approach can be extended to monitor individuals with conditions like Alzheimer's or dementia. A minimal set of 17 features was used to optimize efficiency to recognize eight emotion expressions. At the same time, future work involves expanding the database to include a more diverse range of individuals for improved recognition.

Lohse et al. reviewed the effectiveness of VR therapy in stroke survivors using custom-built virtual environments (VE) and commercially available gaming systems (CG). Out of the twenty-six studies included, the results showed that VR therapy was significantly more effective than conventional therapy in improving body function and activity outcomes. No significant differences were found between VE and CG interventions for these outcomes. However, participation outcomes were mainly derived from VE studies, indicating moderate improvements in VR rehabilitation compared to conventional therapy [2].

Thomas et al. present a modified smart wheelchair to support individuals with restricted mobility. It is based on a commercially available manual wheelchair but includes additional features such as a stereo depth camera, LIDAR, DC gear motor, joystick, and processors. Users have the flexibility to control the wheelchair manually using a joystick or hands-free through a specially developed application. The wheelchair has a Jetson processor, enabling it to navigate autonomously within indoor environments and reach the desired destination. While initially designed for hospitals, this affordable and lightweight wheelchair has the potential to be used in various indoor settings. The project also provides open-source design files, allowing for replication and accessibility to a broader audience [18].

A study by Ivan Yong-Sing, Lau et al. [19] proposed a knee osteoarthritis severity diagnostics system for gait analysis. The researchers collaborated with Sibu Hospital and KPJ Sibu Specialist Hospital to collect subject data. The study utilized the law of cosine and dot cross product as primary measures to analyze gait parameters of the knee, ankle, and hip. The proposed system successfully captures subject movement using the Microsoft Kinect v2 sensor and demonstrates an analysis algorithm for various gait parameters. Future work includes enhancing the system with machine learning algorithms to determine the severity grade of knee osteoarthritis.

The utilization of Kinect extends to supporting the diagnosis and treatment of Scoliosis [20]. For Alzheimer's disease detection, a method is proposed that utilizes a Kinect V.2 camera and machine learning to evaluate the Timed Up and Go (TUG) test and differentiate individuals with Alzheimer's disease (AD) from healthy controls (HC) [21]. Data on joint positions were gathered from both HC and AD participants, and features were extracted from various TUG subtasks. Significant features were identified through statistical analysis and machine learning that could effectively

distinguish AD from HC. This approach demonstrates promise as an accessible and convenient tool for early-stage AD assessment.

As part of a system that enables real-time supervision and monitoring of patients in Intensive Care Units (ICU) within hospital settings, the Kinect sensor is equipped with a range of sensors that track patients' movements, recognize faces and process speech. These sensors do not require physical contact with the patients and instead utilize a Natural User Interface to detect skeletal movements [22].

Limin and Peiyi proposed a gesture recognition method based on LabVIEW and utilizing the 3D somatosensory camera of Kinect for controlling a medical service robot. It tracks human skeleton points, captures real-time human actions, and identifies different body actions on the LabVIEW platform. The Kinect sensor comprises various components such as a primary camera, infrared emitter, infrared depth image camera, rotatable support, and matrix microphone. It captures depth images, colour images, and sound to identify human body movements. The method enables effective human–computer interaction and provides convenience for assistant doctors in completing patient rehabilitation and nursing tasks [23].

In 2022, Amira Gaber et al. developed a system for evaluating facial paralysis (FP), including assessment and classification [24]. The system uses the Kinect V2 sensor to extract real-time facial animation units (FAUs) and employs artificial intelligence and machine learning techniques for classification. A dataset of 375 records from 13 FP patients and 1650 records from 50 control subjects was used to classify seven FP categories, including three severity levels for each side of the face. An ensemble learning classifier based on SVMs was developed to achieve high prediction accuracy. The study demonstrates the effectiveness of using FAUs acquired by the Kinect sensor for classifying FP and highlights the advantages of the developed system over existing grading scales.

Zaid A. Munder and Jiaofei Zhong explore using of a mobile robot and Kinect sensor to develop an intelligent system capable of monitoring and detecting hazardous events such as falls [25]. The Kinect sensor, integrated with the mobile robot, is used to track and detect when a person falls. When a fall is detected, the system sends SMS notifications and makes emergency calls using a mobile phone installed on the robot. The use of depth cameras, such as the Kinect sensor, offers advantages over traditional RGB cameras, providing improved performance, lower costs, and resilience to changes in lighting conditions.

In [26], using Kinect extends to assist blind individuals in navigation by using real-time depth data. The system can detect various environmental patterns, such as obstacles, walls, and stairs. These tasks pose challenges when relying on direct operation by a clinician. However, limitations exist regarding the Kinect sensor's portability and its challenges when capturing depth information in outdoor environments exposed to sunlight or water.

A Microsoft Kinect sensor system has been suggested for exercise control in physiotherapy patients [27]. It adeptly analyzes joint orientations and evaluates muscle group forces, as evidenced by successful experimental outcomes. Moreover, the method holds the potential for exercise control and monitoring applications. The study emphasizes joint range of motion detection and muscle group load analysis,

employing the dot product technique for precise angle calculation. However, the system has not been entailed with a 3D virtual instructor, integrating voice commands and visual movements.

Another method proposes a personal coaching system using a Microsoft Kinect Sensor to monitor and ensure the correct execution of physiotherapy exercises at home [28]. The system includes an exercise-detection algorithm, a recognition module, and a user-friendly web interface. The results demonstrate the successful use of the system for remote physiotherapy exercises, particularly valuable during the COVID-19 pandemic. The Kinect Sensor, known for its movement and speech recognition capabilities, tracks joint coordinates and offers static and dynamic exercises. The system shows promise for remote physiotherapy, and future work aims to integrate it into more advanced virtual coaching systems involving social robots and Ambient Assisted Living environments.

An approach included a system that utilizes gesture recognition of upper body limb movements using the Kinect sensor [29]. The system drives a mobile robotic arm as a prototype. It can be extended to trigger various services for elderly and wheelchair users to help them in their daily activities. The Kinect sensor captures 3D skeleton positions, which are processed by the software tool and converted into data for the microcontroller.

Table 1 showcases additional applications of Kinect sensors within the biomedical and healthcare environment.

Several studies have demonstrated the feasibility and effectiveness of Kinect sensors in medical diagnostic tasks. For instance, Sooklal et al. (2014) explored the detection of tremors using Kinect, providing a potential non-invasive tool for diagnosing conditions like Parkinson's disease [6]. Similarly, Stone et al. (2014) developed a Kinect-based system that generates health alerts from gait measurements, enabling early detection of health changes in gait patterns [31]. These applications showcase the potential of Kinect technology to aid in the assessment and monitoring of various medical conditions.

In the context of Parkinson's disease, Amini Maghsoud Bigy et al. (2015) developed a real-time monitoring system using Kinect to detect Freezing of Gait, tremor, and fall incidents [41]. Their work highlights the potential of Kinect sensors to track joint positions and movements for symptom detection and monitoring. Ren et al. (2019) introduced a multivariate analysis method using Kinect data for Parkinson's disease assessment, demonstrating accuracy in classifying different impairment levels [40]. These studies exemplify the utility of Kinect sensors in advancing the field of medical diagnostics, enabling non-invasive and remote monitoring of various symptoms and conditions.

Kinect technology has shown promise in the realm of rehabilitation and physiotherapy, providing innovative solutions for exercise guidance and assessment. Vogiatzaki et al. (2014) developed a game-based tele-rehabilitation system using Kinect, aiming to enhance stroke patient rehabilitation remotely [32]. Similarly, Saratean et al. (2020) developed a remote physiotherapy application utilizing Kinect, which enabled exercise detection and guidance for patients while allowing physiotherapists to track progress through a web interface [28]. These studies emphasize

Table 1 Applications of Kinect sensor within biomedical and healthcare

Application fields	Year	Authors	Purpose	Method Used	Problems Solved	Object	Results	Weaknesses
Healthcare	2018	Gavrilova et al. [30]	Explore applications of Kinect sensors for gesture and activity recognition in various contexts	Utilized Kinect's capabilities to capture gesture and activity data	Addressed the potential of Kinect sensors for various applications in healthcare, consumer electronics, and cognitive systems	Individuals performing gestures and activities	Demonstrated Kinect's potential in applications ranging from smart homes to biometric surveillance	Acknowledged challenges of context-based recognition and data privacy
Medical diagnostic	2014	Sooklal et al. [6]	Detect tremors using Kinect	Simulation of tremors, Kinect data capture	Detect Parkinsonian, postural, voice tremors	Accuracy in detecting head tremors, sensitivity to background movement	Kinect sensor, tremors	Potential for home monitoring and therapy
Medical diagnostic	2014	Stone et al. [31]	Generate health alerts from gait measurements	Kinect-based gait measurement system	Early detection of health changes in gait patterns	Modeling uncertainty in gait parameter estimates	Gait measurement system	Alerts for potential health issues
Rehabilitation	2014	Vogiatzaki et al. [32]	Develop game-based tele-rehabilitation system	Kinect, virtual reality, cloud-based system	Game-based rehabilitation for stroke patients	Remote exercise monitoring, bio-feedback	Tele-rehabilitation system	Enhance stroke patient rehabilitation

(continued)

Table 1 (continued)

Application fields	Year	Authors	Purpose	Method Used	Problems Solved	Object	Results	Weaknesses
Medical diagnostic, rehabilitation	2015	Pineda-Lopez et al. [33]	Develop gait analysis system using Kinect	Kinect sensor, database system	Gait analysis and visualization	Diagnosing walking problems, rehabilitation	Human body movement analysis	Gait cycle classification
Rehabilitation	2016	Pathirana et al. [34]	Develop real-time tele-rehabilitation system	Kinect, multiple devices, Kalman filter	Accurate joint position estimation	Robust bio-kinematic tracking	Tele-rehabilitation system	Improved movement tracking
Medical diagnostic	2017	Tanaka and Sogabe [35]	Measure and classify leg shapes	Kinect sensor, algorithm	Automated leg shape measurement	Classify genu valgum and genu varum	Leg shapes, Kinect sensor	Accurate leg shape classification
Medical diagnostic	2013	Shao et al. [36]	Introduce Kinect-like sensors for computer vision applications	Reviewed applications of Kinect-like sensors for computer vision	Explored the potential of Kinect-like sensors in computer vision applications	RGB-D sensor applications in computer vision	Highlighted advancements in object recognition, 3D mapping, and scene classification	Addressed challenges in depth computation and fusion
Medical diagnostic	2017	Jawaid and Qureshi [3]	Review of Kinect-based medical imaging applications	Explored applications of Kinect for medical imaging tasks	Highlighted Kinect's use in medical imaging for diagnostics and treatments	Medical imaging using Kinect	Demonstrated Kinect's applications in medical diagnostics	Addressed challenges of noise, accuracy, and sensor fusion

(continued)

Table 1 (continued)

Application fields	Year	Authors	Purpose	Method Used	Problems Solved	Object	Results	Weaknesses
Medical diagnostic	2018	Eltoukhy et al. [37]	Validate Kinect v2 sensor for assessing balance in both young and elderly populations	Utilized Kinect v2 to capture balance-related data during static and dynamic tasks	Addressed limitations in assessing balance accurately and conveniently	Young and elderly participants performing balance tasks	Kinect v2 provided valid measurements of balance-related parameters for both age groups	Limitations included small sample size and challenges in temporal variations
Medical diagnostic	2019	Saraguro et al. [38]	Develop a system to analyze hand movements in PD patients using Kinect	Utilized Kinect V2 sensor, tracked hand movements for PD evaluation	Provided non-invasive tool for evaluating PD progression via finger tapping tests	Hand movements of PD patients	Detected lower tapping velocity in upper right extremities of PD patients	Variation in medication status, sensitivity to disease severity
Medical diagnostic	2019	Dehbandi et al. [39]	Explore the feasibility of using Kinect 2 data for quantifying upper limb behavior in motor disorders	Developed a method to quantify joint motion data using Kinect 2	Addressed limitations in accurately capturing motor symptoms and movement patterns in motor disorders	Participants with and without motor disorders performing sit-to-stand-to-sit motion	Proposed method showed exceptional classification performance for distinguishing healthy and motor disorder participants	Challenges included computational demands of the method and broader assessment of motor symptoms

(continued)

Table 1 (continued)

Application fields	Year	Authors	Purpose	Method Used	Problems Solved	Object	Results	Weaknesses
Training	2017	Limin and Peiyi [23]	Develop a gesture recognition method for controlling a medical service robot using Kinect	Kinect technology, LabVIEW software	Enhancing human–computer interaction in medical service robots	Medical service robot	Successful human–computer interaction for controlling medical service robots in healthcare settings	Acknowledges limitations of Kinect positioning and potential accuracy issues due to environmental factors
Medical diagnostic	2019	Ren et al. [40]	Introduce a multivariate analysis method using Kinect data for Parkinson's disease assessment	Developed a method to analyze joint motion data using Kinect for Parkinson's disease assessment	Addressed limitations in accurately capturing motor symptoms in Parkinson's disease	Healthy participants and Parkinson's disease patients performing sit-to-stand-to-sit motion	Proposed method demonstrated accuracy in classifying Parkinson's disease impairment levels	Challenges included computational demands and limited assessment of motor symptoms

(continued)

Table 1 (continued)

Application fields	Year	Authors	Purpose	Method Used	Problems Solved	Object	Results	Weaknesses
Monitoring, medical diagnostic	2015	Amini Maghsoud Bigy et al. [41]	Develop a Kinect-based system for real-time monitoring and detection of Freezing of Gait, tremor, and fall incidents in Parkinson's disease patients	Utilized Kinect sensor to track joint positions and movements for symptom detection	Addressed challenges of detecting critical incidents in Parkinson's disease patients	Parkinson's disease patients exhibiting various symptoms	Demonstrated Kinect's potential for detecting FOG, tremor, and falls with promising accuracy	Limitations included simulated subjects and potential tracking errors
Physiotherapy	2017	Cubukcu and Yuzgec [42]	Develop a Kinect-based physiotherapy application for patients with shoulder joint, muscle, and tendon damage	Created a physiotherapy application using Kinect for exercise guidance and assessment	Addressed difficulties in accessing physiotherapy centers and performing exercises correctly	Patients with shoulder joint, muscle, and tendon damage	Demonstrated the potential of Kinect-based application for remote physiotherapy guidance	Limitations included focusing on specific conditions and lack of interactive communication
Physiotherapy	2020	Saratean et al. [28]	Develop a system for remote physiotherapy using Kinect	Utilized Kinect Sensor for exercise detection, web interface for physiotherapists	Supported remote physiotherapy through exercise guidance and tracking	Physiotherapy exercises	Detected exercise correctness and guided patients remotely	Limitation in tracking obscured joints, Kinect's field of view limitation

(continued)

Table 1 (continued)

Application fields	Year	Authors	Purpose	Method Used	Problems Solved	Object	Results	Weaknesses
Physiotherapy	2020	Huang et al. [43]	Develop a Kinect-based self-help training system for nursing students to learn patient transfer skills	Utilized Kinect RGB-D sensors to measure trainee and patient postures for skill evaluation	Addressed the limitations of traditional nursing education for patient transfer skills	Nursing students learning patient transfer skills	Demonstrated accuracy and effectiveness of the system for improving patient transfer skills	Addressed tracking errors and potential challenges in evaluation
Rehabilitation	2014	Gauthier and Cretu [44]	Develop a Kinect-based system for monitoring and assessing exercise performance in home-based settings	Utilized Kinect to track joint movements and quantify exercise performance	Addressed limitations in monitoring exercise performance and adherence in home environments	Individuals performing physical exercise at home	Demonstrated the feasibility of using Kinect for quantifying exercise movements and providing feedback	Limitations included simplified joint model and potential measurement errors
Rehabilitation	2015	Lai et al. [45]	Develop a Kinect-based system for virtual rehabilitation to train balance in stroke patients	Utilized Kinect V1 to create a virtual rehabilitation system	Addressed limitations of in-person rehabilitation for stroke patients by providing a home-based training solution	Stroke patients performing balance training exercises	Significant improvements in balance-related parameters for stroke patients using the Kinect-based system	Challenges included further exploration of exercises and ensuring consistent participation

(continued)

Table 1 (continued)

Application fields	Year	Authors	Purpose	Method Used	Problems Solved	Object	Results	Weaknesses
Medical education	2015	Pauly et al. [46]	Improve surgical scene understanding	Augmented reality using Kinect and X-ray data fusion	Enhance spatial understanding in orthopedic surgery	Augmented reality setup	Improved surgical scene understanding, depth perception enhancement	Challenges in object recognition using color and depth cues
Rehabilitation	2017	Park et al. [47]	Assess VR training effects on motor recovery	Virtual reality training with Xbox Kinect	Improve motor function and recovery in stroke survivors	Stroke survivors	VR training improved motor function, efficient instrument movement	Discrepancy in intervention time, Kinect recognition errors
Medical education	2018	Feng et al. [48]	Evaluate Virtual Pointer in laparoscopic training	Virtual Pointer system using Kinect for guidance	Enhance professional vision in laparoscopic surgical training	Surgical trainees	Virtual Pointer improved perception and economy of movement	No significant reduction in errors or time to completion
Medical services	2014	Kim et al. [49]	Compare Kinect hand tracking to mechanical control	Kinect hand tracking technology for surgical robots	Explore feasibility of Kinect-based control for surgical robots	Experienced surgeons	Kinect hand tracking is promising but needs improvement	Latency and static noise affecting performance
Rehabilitation, monitoring	2014	Saiyi Li et al. [50]	Improve Kinect-based rehabilitation monitoring	Multi-Kinect fusion for joint position measurement	Enhance accuracy of joint position measurement in rehab monitoring	Rehabilitation patients	Improved joint position measurement accuracy	Noise in Kinect data, limited dataset size

(continued)

Table 1 (continued)

Application fields	Year	Authors	Purpose	Method Used	Problems Solved	Object	Results	Weaknesses
Medical services	2017	Xiao et al. [51]	Develop Kinect-based navigation system	Optical navigation using Kinect camera	Enhance accuracy of needle insertion in abdominal interventions	Phantom and animal experiments	Accurate needle placement, better accuracy with Kinect V2	Complexity of workflow, need for respiratory compensation
Medical education	2022	Fuchs et al. [52]	Develop evaluation system for endoscopic training	Multimodal feature combination using Kinect and EMG	Evaluate learning benefits of stereoscopic views in training	Surgical trainees	Validated simulator-independent endoscopic skill assessment	Small dataset size, need for more extensive training

the potential of Kinect-based systems to support rehabilitation and physiotherapy efforts by offering remote guidance and monitoring capabilities.

Lai et al. (2015) developed a Kinect-based virtual rehabilitation system to train balance in stroke patients, providing an alternative to in-person rehabilitation [45]. The findings suggest significant improvements in balance-related parameters for stroke patients using the Kinect-based system. Additionally, Cubukcu and Yuzgec (2017) created a Kinect-based physiotherapy application for patients with shoulder joint, muscle, and tendon damage, offering exercise guidance and assessment outside of traditional healthcare settings [42]. These studies highlight the potential of Kinect technology to revolutionize the field of rehabilitation, making exercise guidance and assessment more accessible and convenient for patients.

In the realm of surgical training and education, Kinect technology has been explored to enhance surgical skills and professional vision. Feng et al. (2018) evaluated the effectiveness of the Virtual Pointer system in improving laparoscopic surgical training by using Kinect for visual guidance [48]. The results indicated improved perception and economy of movement, enhancing trainees' adoption of professional vision. Moreover, Kim et al. (2014) investigated the feasibility of using Kinect-based hand tracking technology for controlling surgical robots, showcasing its potential to revolutionize surgical control methods [49]. These studies illustrate the potential of Kinect technology to enhance surgical training and control systems, offering innovative solutions to improve skill development and performance.

These applications above show that VR has positively impacted the biomedical field and healthcare environment. The application of VR will continue to develop and experience further improvements to help improve diagnosis, therapy, health education, and many more.

3.1 Challenges and Opportunities in the Use of VR and Kinect Sensor

While the integration of VR and Kinect technology holds promise in healthcare, there are challenges to consider. Accuracy [41], potential tracking errors [3, 44], and computational demands [40] are among the limitations identified in various studies. Additionally, the adoption of these technologies in clinical settings requires careful consideration of patient safety, data privacy [30], and effective integration into existing healthcare workflow.

The algorithm's performance might vary based on individual characteristics, and further refinements and real-time testing are needed to validate its reliability. The study also acknowledged the possibility of false positives or missed alerts and emphasized the importance of parameter tuning by clinicians [31, 53].

The accurate and reliable capture of data stands as a central challenge in applications reliant on both Virtual Reality (VR) and Kinect-based technology. The studies underscore the critical demand for precision in tracking movements, particularly

within clinical contexts where precise data is imperative for diagnostic and therapeutic purposes. The calibration and setup of Kinect sensors represent essential factors in attaining accurate data [54]. Ensuring the precise positioning and orientation of these sensors can prove challenging, particularly when dealing with home-based environments. Additionally, the presence of environmental variables, occlusions, and noise within depth data [6] can significantly impact the accuracy of tracking and recognition algorithms. The resolution of these issues is of paramount importance to establish dependable and consistent outcomes.

The development of robust algorithms to facilitate gesture recognition, movement tracking, and pose estimation [53] introduces a complex undertaking. This necessitates careful consideration of numerous factors, including the selection of appropriate machine learning models and feature extraction techniques. Furthermore, the diverse nature of human anatomy and movements across individuals accentuates the need for algorithms and systems that can adeptly adapt to accommodate these differences. This ensures the technology's effectiveness across a wide spectrum of users.

Real-time processing [32], especially in applications like tele-rehabilitation and human–robot interaction, introduces challenges due to the requisite low latency and immediate feedback. The integration of VR and Kinect-based systems into established clinical practices mandates validation, regulatory compliance, and alignment with healthcare standards. Clinicians seek assurance regarding the technology's precision, safety, and efficacy, thereby emphasizing the need for comprehensive validation and adherence to industry regulations.

While using virtual reality (VR) and the Kinect sensor has challenges in various fields, it also brings numerous opportunities. VR and Kinect technologies present a remarkable avenue for remote monitoring and rehabilitation, enabling patients to seamlessly engage in therapy from the comfort of their homes. This holds the potential to ameliorate patient compliance and alleviate the strain on healthcare facilities. Integrating gamification and interactive exercises into rehabilitation programs imparts an engaging and motivating dimension. The immersive qualities of VR, coupled with the natural interaction facilitated by Kinect sensors, synergistically contribute to heightened patient adherence and participation rates.

Beyond patient engagement, the amalgamation of VR and Kinect-based systems yields a trove of movement data that can be systematically analyzed to yield insights into patient progress, movement patterns, and avenues for potential enhancement. This data-centric approach paves the way for personalized rehabilitation regimens that cater to individual needs.

A distinctive advantage of Kinect sensors lies in their cost-effectiveness as a depth-sensing solution, particularly when juxtaposed with more intricate motion capture systems. The affordability of Kinect sensors extends opportunities for broader adoption, particularly in settings with resource constraints. The meticulous tracking and detailed movement analysis facilitated by Kinect sensors hold profound diagnostic potential for identifying movement disorders and evaluating the efficacy of treatments. Moreover, VR-based diagnostic tools, enriched by Kinect data, proffer clinicians with invaluable insights to inform their decision-making.

VR and Kinect systems exhibit remarkable adaptability, allowing interventions to be tailored to individual progress and capabilities. This adaptiveness, in conjunction with the synergy between VR and Kinect, extends beyond rehabilitation to enrich human–robot interaction. The fusion of these technologies empowers robots to comprehend and respond to human gestures and movements with heightened intuition, fostering a more seamless interaction between humans and machines.

The collective studies underscore the strides made in the field, yet concurrently emphasize the expansive realm for further research and innovation. The path forward beckons the development of novel algorithms, the refinement of existing technologies, and the exploration of untapped applications. As these technologies mature, the marriage of VR and Kinect is poised to revolutionize healthcare, offering solutions that transcend physical boundaries and cater to individual needs with unprecedented precision.

While the studies presented valuable insights, there are several limitations to consider. First, the reviewed studies relied on relatively small sample sizes, limiting the generalizability of the findings. The studies predominantly focused on specific medical conditions, potentially excluding broader applications. Furthermore, the challenges related to VR and Kinect technology, such as hardware limitations and user comfort, were not always extensively explored within the reviewed studies.

4 Conclusion

Virtual reality (VR) has significantly changed various industries, including biomedical and health. By immersing individuals in computer-generated environments, VR offers interactive and immersive experiences, and its applications extend beyond being a medium or advanced user interface. VR has practical solutions to real-world challenges, and its effectiveness relies on immersion, interaction, and imagination. The history of VR spans over 40 years, with significant advancements in hardware and software components. The Kinect sensor, initially developed for gaming, has found applications in biomedical and health fields, such as diagnostic tasks, monitoring of various medical conditions, rehabilitation and physiotherapy, surgical training and education. While challenges exist, VR and the Kinect sensor provide opportunities for personalized interventions, remote healthcare delivery, training, collaboration, and improved patient engagement. As technology evolves, the potential of VR and the Kinect sensor in healthcare is expected to grow, with a focus on managing costs, enhancing hardware capabilities, ensuring data security, and refining user experiences.

References

1. Pensieri C, Pennacchini M (2016) Virtual reality in medicine. https://doi.org/10.1007/978-3-319-22041-3_14
2. Lohse KR, Hilderman CGE, Cheung KL, Tatla S, Van der Loos HFM (2014) Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. PLoS ONE 9:e93318. https://doi.org/10.1371/journal.pone.0093318
3. Jawaid I, Qureshi JK (2017) Advancements in medical imaging through Kinect: a review. In: 2017 international symposium on wireless systems and networks (ISWSN). IEEE, pp 1–5. https://doi.org/10.1109/ISWSN.2017.8250032
4. Dargar S, Nunno A, Sankaranarayanan G, De S (2013) Microsoft Kinect based head tracking for life size collaborative surgical simulation environments (LS-CollaSSLE). Stud Health Technol Inform 184:109–113
5. Palter VN, Grantcharov TP (2010) Virtual reality in surgical skills training. Surg Clin North Am 90:605–617. https://doi.org/10.1016/j.suc.2010.02.005
6. Sooklal S, Mohan P, Teelucksingh S (2014) Using the Kinect for detecting tremors: challenges and opportunities. In: IEEE-EMBS international conference on biomedical and health informatics (BHI). IEEE, pp 768–771. https://doi.org/10.1109/BHI.2014.6864477
7. Suarez J, Murphy RR (2012) Hand gesture recognition with depth images: a review. In: 2012 IEEE RO-MAN: the 21st IEEE international symposium on robot and human interactive communication. IEEE, pp 411–417. https://doi.org/10.1109/ROMAN.2012.6343787
8. Burdea GC, Coiffet P (2003) Virtual reality technology. Wiley
9. Burdea GC (1999) Invited review: the synergy between virtual reality and robotics. IEEE Trans Robot Autom 15:400–410. https://doi.org/10.1109/70.768174
10. Zhang Z (2012) Microsoft Kinect sensor and its effect. IEEE Multimed 19:4–10. https://doi.org/10.1109/MMUL.2012.24
11. Yang L, Zhang L, Dong H, Alelaiwi A, El Saddik A (2015) Evaluating and improving the depth accuracy of Kinect for Windows v2. IEEE Sens J 15:4275–4285. https://doi.org/10.1109/JSEN.2015.2416651
12. El-laithy RA, Huang J, Yeh M (2012) Study on the use of Microsoft Kinect for robotics applications. In: Proceedings of the 2012 IEEE/ION position, location and navigation symposium. IEEE, pp 1280–1288. https://doi.org/10.1109/PLANS.2012.6236985
13. Correa DSO, Sciotti DF, Prado MG, Sales DO, Wolf DF, Osório FS (2012) Mobile robots navigation in indoor environments using Kinect sensor. In: Proceedings - 2012 2nd Brazilian conference on critical embedded systems, CBSEC 2012, pp 36–41. https://doi.org/10.1109/CBSEC.2012.18
14. Eric N, Jang J-W (2017) Kinect depth sensor for computer vision applications in autonomous vehicles. In: 2017 ninth international conference on ubiquitous and future networks (ICUFN). IEEE, pp 531–535. https://doi.org/10.1109/ICUFN.2017.7993842
15. Clark RA, Pua Y-H, Oliveira CC, Bower KJ, Thilarajah S, McGaw R, Hasanki K, Mentiplay BF (2015) Reliability and concurrent validity of the Microsoft Xbox One Kinect for assessment of standing balance and postural control. Gait Posture 42:210–213. https://doi.org/10.1016/j.gaitpost.2015.03.005
16. Kassem A, Hamad M, El Moucary C, Nawfal E, Aoun A (2017) MedBed: smart medical bed. In: 2017 Fourth international conference on advances in biomedical engineering (ICABME). IEEE, pp 1–4. https://doi.org/10.1109/ICABME.2017.8167544
17. Alabbasi HA, Moldoveanu F, Moldoveanu A, Shhedi Z (2015) Facial emotion expressions recognition with brain activites using Kinect sensor V2. Int Res J Eng Technol
18. Sola-Thomas E, Baser Sarker MA, Caracciolo MV, Casciotti O, Lloyd CD, Imtiaz MH (2021) Design of a low-cost, lightweight smart wheelchair. In: 2021 IEEE microelectronics design and test symposium, MDTS 2021. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/MDTS52103.2021.9476093

19. Lau IY-S, Chua T-T, Lee WX-P, Wong C-W, Toh T-H, Ting H-Y (2020) Kinect-based knee osteoarthritis gait analysis system. In: 2020 IEEE 2nd international conference on artificial intelligence in engineering and technology (IICAIET). IEEE, pp 1–6. https://doi.org/10.1109/IICAIET49801.2020.9257860
20. Cukovic S, Petruse RE, Meixner G, Buchweitz L (2020) Supporting diagnosis and treatment of scoliosis: using augmented reality to calculate 3D spine models in real-time - ARScoliosis. In: 2020 IEEE international conference on bioinformatics and biomedicine (BIBM). IEEE, pp 1926–1931. https://doi.org/10.1109/BIBM49941.2020.9313200
21. Seifallahi M, Mehraban AH, Galvin JE, Ghoraani B (2022) Alzheimer's disease detection using comprehensive analysis of timed up and go test via Kinect vol 2 Camera and machine learning. IEEE Trans Neural Syst Rehabil Eng 30:1589–1600. https://doi.org/10.1109/TNSRE.2022.3181252
22. Shehabat IM, Al-Hussein N (2018) Deploying internet of things in healthcare: benefits, requirements, challenges and applications. J Commun 574–580. https://doi.org/10.12720/jcm.13.10.574-580
23. Limin M, Peiyi Z (2017) The medical service robot interaction based on Kinect. In: 2017 IEEE international conference on intelligent techniques in control, optimization and signal processing (INCOS). IEEE, pp 1–7. https://doi.org/10.1109/ITCOSP.2017.8303077
24. Gaber A, Taher MF, Wahed MA, Shalaby NM, Gaber S (2022) Classification of facial paralysis based on machine learning techniques. Biomed Eng Online 21:65. https://doi.org/10.1186/s12938-022-01036-0
25. Mundher ZA, Jiaofei Z (2014) A real-time fall detection system in elderly care using mobile robot and Kinect sensor. Int J Mater Mech Manuf 2:133–138. https://doi.org/10.7763/IJMMM.2014.V2.115
26. Filipe V, Fernandes F, Fernandes H, Sousa A, Paredes H, Barroso J (2012) Blind navigation support system based on Microsoft Kinect. Procedia Comput Sci 14:94–101. https://doi.org/10.1016/j.procs.2012.10.011
27. Topuz B, Ozuag E, Akbulut O (2016) Kinect sensor based physiotherapy management. In: 2016 24th signal processing and communication application conference (SIU). IEEE, pp 2209–2212. https://doi.org/10.1109/SIU.2016.7496213
28. Saratean T, Antal M, Pop C, Cioara T, Anghel I, Salomie I (2020) A physiotheraphy coaching system based on Kinect sensor. In: 2020 IEEE 16th international conference on intelligent computer communication and processing (ICCP). IEEE, pp 535–540. https://doi.org/10.1109/ICCP51029.2020.9266178
29. Ababneh M, Shaban H, AlShalabe D, Khader D, Mahameed H, AlQudimat M (2018) Gesture controlled mobile robotic arm for elderly and wheelchair people assistance using Kinect sensor. In: 2018 15th international multi-conference on systems, signals and devices (SSD). IEEE, pp 636–641. https://doi.org/10.1109/SSD.2018.8570547
30. Gavrilova ML, Wang Y, Ahmed F, Polash Paul P (2018) Kinect sensor gesture and activity recognition: new applications for consumer cognitive systems. IEEE Consum Electron Mag 7:88–94. https://doi.org/10.1109/MCE.2017.2755498
31. Stone EE, Skubic M, Back J (2014) Automated health alerts from Kinect-based in-home gait measurements. In: 2014 36th annual international conference of the IEEE engineering in medicine and biology society. IEEE, pp 2961–2964. https://doi.org/10.1109/EMBC.2014.6944244
32. Vogiatzaki E, Gravezas Y, Dalezios N, Biswas D, Cranny A, Ortmann S, Langendorfer P, Lamprinos I, Giannakopoulou G, Achner J, Klemke J, Jost H (2014) Telemedicine system for game-based rehabilitation of stroke patients in the FP7-"StrokeBack" project. In: 2014 European conference on networks and communications (EuCNC). IEEE, pp 1–5. https://doi.org/10.1109/EuCNC.2014.6882688
33. Pineda-Lopez FM, Flores CMJ, Ortiz VGJ, Mosquera YWA (2015) Prototype for the analysis of human body movement with Kinect sensor. In: 2015 20th symposium on signal processing, images and computer vision (STSIVA). IEEE, pp 1–7. https://doi.org/10.1109/STSIVA.2015.7330410

34. Pathirana PN, Li S, Trinh HM, Seneviratne A (2016) Robust real-time bio-kinematic movement tracking using multiple Kinects for tele-rehabilitation. IEEE Trans Industr Electron 63:1822–1833. https://doi.org/10.1109/TIE.2015.2497662
35. Tanaka M, Sogabe A (2017) A measuring system of the legs shape by using the Kinect sensor. In: 2017 56th annual conference of the society of instrument and control engineers of Japan (SICE). IEEE, pp 106–109. https://doi.org/10.23919/SICE.2017.8105558
36. Shao L, Han J, Xu D, Shotton J (2013) Computer vision for RGB-D sensors: Kinect and its applications [special issue intro]. IEEE Trans Cybern 43:1314–1317. https://doi.org/10.1109/TCYB.2013.2276144
37. Eltoukhy MA, Kuenze C, Oh J, Signorile JF (2018) Validation of static and dynamic balance assessment using Microsoft Kinect for young and elderly populations. IEEE J Biomed Health Inform 22:147–153. https://doi.org/10.1109/JBHI.2017.2686330
38. Saraguro W, Barzallo B, Guillermo J, Garcia-Cedeno A, Soto A, Rivas D, Clotet R, Huerta M (2019) Analysis of hand movements in patients with Parkinson's disease using Kinect. In: 2019 IEEE international conference on e-health networking, application and services (HealthCom). IEEE, pp 1–6. https://doi.org/10.1109/HealthCom46333.2019.9009589
39. Dehbandi B, Barachant A, Harary D, Long JD, Tsagaris KZ, Bumanlag SJ, He V, Putrino D (2017) Using data from the Microsoft Kinect 2 to quantify upper limb behavior: a feasibility study. IEEE J Biomed Health Inform 21:1386–1392. https://doi.org/10.1109/JBHI.2016.2606240
40. Ren P, Bosch Bayard JF, Dong L, Chen J, Mao L, Ma D, Sanchez MA, Morejon DM, Bringas ML, Yao D, Jahanshahi M, Valdes-Sosa PA (2020) Multivariate analysis of joint motion data by Kinect: application to Parkinson's disease. IEEE Trans Neural Syst Rehabil Eng 28:181–190. https://doi.org/10.1109/TNSRE.2019.2953707
41. Amini Maghsoud Bigy A, Banitsas K, Badii A, Cosmas J (2015) Recognition of postures and Freezing of Gait in Parkinson's disease patients using Microsoft Kinect sensor. In: 2015 7th international IEEE/EMBS conference on neural engineering (NER). IEEE, pp 731–734. https://doi.org/10.1109/NER.2015.7146727
42. Cubukcu B, Yuzgec U (2017) A physiotherapy application with MS Kinect for patients with shoulder joint, muscle and tendon damage. In: 2017 9th international conference on computational intelligence and communication networks (CICN). IEEE, pp 225–228. https://doi.org/10.1109/CICN.2017.8319390
43. Huang Z, Nagata A, Kanai-Pak M, Maeda J, Kitajima Y, Nakamura M, Aida K, Kuwahara N, Ogata T, Ota J (2014) Self-help training system for nursing students to learn patient transfer skills. IEEE Trans Learn Technol 7:319–332. https://doi.org/10.1109/TLT.2014.2331252
44. Gauthier S, Cretu A-M (2014) Human movement quantification using Kinect for in-home physical exercise monitoring. In: 2014 IEEE international conference on computational intelligence and virtual environments for measurement systems and applications (CIVEMSA). IEEE, pp 6–11. https://doi.org/10.1109/CIVEMSA.2014.6841430
45. Lai C-L, Huang Y-L, Liao T-K, Tseng C-M, Chen Y-F, Erdenetsogt D (2015) A Microsoft Kinect-based virtual rehabilitation system to train balance ability for stroke patients. In: 2015 international conference on cyberworlds (CW). IEEE, pp 54–60. https://doi.org/10.1109/CW.2015.44
46. Pauly O, Diotte B, Fallavollita P, Weidert S, Euler E, Navab N (2015) Machine learning-based augmented reality for improved surgical scene understanding. Comput Med Imaging Graph 41:55–60. https://doi.org/10.1016/j.compmedimag.2014.06.007
47. Park D-S, Lee D-G, Lee K, Lee G (2017) Effects of virtual reality training using Xbox Kinect on motor function in stroke survivors: a preliminary study. J Stroke Cerebrovasc Dis 26:2313–2319. https://doi.org/10.1016/j.jstrokecerebrovasdis.2017.05.019
48. Feng Y, McGowan H, Semsar A, Zahiri HR, George IM, Turner T, Park A, Kleinsmith A, Mentis HM (2018) A virtual pointer to support the adoption of professional vision in laparoscopic training. Int J Comput Assist Radiol Surg 13:1463–1472. https://doi.org/10.1007/s11548-018-1792-9

49. Kim Y, Leonard S, Shademan A, Krieger A, Kim PCW (2014) Kinect technology for hand tracking control of surgical robots: technical and surgical skill comparison to current robotic masters. Surg Endosc 28:1993–2000. https://doi.org/10.1007/s00464-013-3383-8
50. Li S, Pathirana PN, Caelli T (2014) Multi-Kinect skeleton fusion for physical rehabilitation monitoring. In: 2014 36th annual international conference of the IEEE engineering in medicine and biology society. IEEE, pp 5060–5063. https://doi.org/10.1109/EMBC.2014.6944762
51. Xiao D, Luo H, Jia F, Zhang Y, Li Y, Guo X, Cai W, Fang C, Fan Y, Zheng H, Hu Q (2016) A Kinect^{TM} camera based navigation system for percutaneous abdominal puncture. Phys Med Biol 61:5687–5705. https://doi.org/10.1088/0031-9155/61/15/5687
52. Fuchs R, Van Praet KM, Bieck R, Kempfert J, Holzhey D, Kofler M, Borger MA, Jacobs S, Falk V, Neumuth T (2022) A system for real-time multivariate feature combination of endoscopic mitral valve simulator training data. Int J Comput Assist Radiol Surg 17:1619–1631. https://doi.org/10.1007/s11548-022-02588-1
53. Tellaeche A, Maurtua I (2014) 6DOF pose estimation of objects for robotic manipulation. A review of different options. In: Proceedings of the 2014 IEEE emerging technology and factory automation (ETFA). IEEE, pp 1–8. https://doi.org/10.1109/ETFA.2014.7005077
54. Bt Ismail NH, Basah SNB (2015) The applications of Microsoft Kinect for human motion capture and analysis : a review. In: 2015 2nd international conference on biomedical engineering (ICoBE). IEEE, pp 1–4. https://doi.org/10.1109/ICoBE.2015.7235894

Automated Taekwondo Kick Classification Using SVM and IMU Sensor on Arduino Nano 33 BLE

Qoriina Dwi Amalia, Azhar Agustian Gunawan, Grachia Salsabila Yulian, Achmad Rizal, and Istiqomah

Abstract Since practically all activities were impeded by the COVID-19 epidemic, they were conducted as autonomously as possible at home, or what is known as Work from Home (WFH). Taekwondo activities are among those that cannot be performed as WFH. The COVID-19 epidemic disrupted regular Taekwondo training, necessitating autonomous practice at home. However, without a trainer's presence, technical errors in Taekwondo kicks could occur. The research presents an automated system utilizing an IMU sensor and SVM for Taekwondo kick classification, empowering athletes to improve their movements independently. Type of kicks that can be classified are Eolgol Ap Chagi, Momtong Ap Chagi, Eolgol Dollyo Chagi, Momtong Dollyo Chagi, and Dwi Chagi. Because it is a basic kick that must be mastered by Taekwondo Athletes. When using tools, taekwondo athletes can move more quickly, thanks to direct implementation on compact devices. Consequently, a simple machine-learning model with the fewest input characteristics is required. On the Arduino Nano 33 BLE, the LSM9DS IMU sensor was used to collect the data. The dataset goes through a cleansing procedure before being labelled and trained, which comes before pre-processing. Three options that may have been employed are SVM, RBF, and DT. In this study, the micromlgen library will be used. Consequently, this work employs a Support Vector Machine (SVM) methodology. Mean, median, max, min, and variance are the five features used in the pre-processing technique. The median and variance properties are used to get an accuracy of 99.35%. The experimental findings demonstrate that the SVM algorithm successfully categorizes the different kinds of taekwondo kicks. The developed technology serves as a valuable tool for Taekwondo athletes, providing a means to enhance their skills through self-guided practice during situations like the COVID-19 pandemic.

Keywords Taekwondo · Machine learning · Feature selection · Time domain

Q. D. Amalia · A. A. Gunawan · G. S. Yulian · A. Rizal (✉) · Istiqomah
School of Electrical Engineering, Telkom University, Bandung 40257, Indonesia
e-mail: achmadrizal@telkomuniversity.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_3

1 Introduction

One of the most famous martial arts branches originating from Korea is Taekwondo. Taekwondo itself is divided into three basic words, including "tae" which means kick technique, "kwon" which means punch technique, and "do" which means martial art with bare feet and hand techniques. Taekwondo is very useful in training the motor components of the body, such as coordination, speed, balance, power, endurance, posture, and flexibility [1]. In Taekwondo, there is one basic category called poomsae. Poomsae or stance attack is a series of basic self-defense techniques or attacks that aim to fight [2]. One of the moves that exist is the kick technique. Kick technique involves various aspects that influence its success, including speed and power in producing optimal kicks [2]. Much research has been done on the Recognition of Human Activity or Human Activity Recognition (HAR) [3]. However, in the context of Taekwondo, there still needs to be more research [4].

Research related to the classification of human movement generally uses Inertial Measurement Unit (IMU) sensors to retrieve data [3]. In the previous study, a classification test of two types of kicks was carried out, namely Ap Chagi and Dollyo Chagi, using the K-Nearest Neighbors (K-NN) and Support Vector Machine (SVM) methods [5]. Regarding accuracy calculations, it is proven that the K-NN method is superior to SVM. This research was implemented on the Raspberry Pi Zero [6]. However, the weakness is in the size of the device, which is still large enough to make it difficult for Taekwondo athletes to perform kick movements. Using the SVM method on machine learning which will simplify the process of classifying the types of direct kicks on the tool.

This follow-up study aims to classify five types of Taekwondo kicks using smaller devices so athletes can feel comfortable using them. The device used in this research is Arduino Nano 33 BLE. To run machine learning on these devices, the choice of classification method cannot be arbitrary because the micromlgen library is needed for its implementation [7]. Several classification algorithms that can be implemented include Decision Tree, Random Forest, and Support Vector Machine. This research will use the LSM9DS1 sensor to retrieve the dataset. Of all the methods tested, only the best method will be used in the implementation [3, 8].

During the COVID-19 pandemic, Taekwondo practice, usually done routinely and together, had to stop completely. Not only practice but even championships with the kyukpa, poomsae, and kyorugi categories were also affected. In India, championships have been held by asking participants to make videos on YouTube channels as an alternative [5, 6]. However, this approach is somewhat less effective. Therefore, this study proposes the development of a tool that can help Taekwondo athletes in training and championships in the Poomsae category [2]. The objectives in this study include:

a. Designing a device that can classify the types of kicks in Taekwondo.
b. This tool is designed to classify five types of kicks in Taekwondo, namely Eolgol Ap Chagi, Momtong Ap Chagi, Eolgol Dollyo Chagi, Momtong Dollyo Chagi, and Dwi Chagi.
c. The device can communicate wirelessly.

The development of this tool is essential as it can assist Taekwondo athletes during the pandemic and beyond, enabling them to continue their training and participate in championships more effectively.

2 Materials and Method

This research method includes taking datasets using accelerometer sensors, followed by normalization processes and analysis using machine learning algorithms for classification purposes. The output of the developed model is then analyzed based on accuracy, precision, and recall. Statistical features have an important role in achieving high accuracy. All of these methodologies are explained in detail in Fig. 1.

In Fig. 1, the IMU sensor will obtain five types of taekwondo kick movements in the form of Momtong Ap Chagi, Eolgol Ap Chagi, Momtong Dollyo Chagi, Eolgol Dollyo Chagi, and Dwi Chagi. These five datasets are then integrated and go through a pre-processing process. Here, a feature selection trial will be carried out. Then, the data goes through a training process by selecting parameters in the model using GridsearchCV and will then be evaluated to determine how accurate the predictions.

2.1 Dataset Retrieval

The initial stage involves taking a dataset by collecting raw data using the LSM9DS1 accelerometer sensor. The IMU sensor is placed on the subject's ankle to detect the type of movement being performed as displayed in Fig. 2. This data was collected from five Taekwondo athletes with different body posture variations, with height 167 $\pm$ 9.8 cm and weight 63 $\pm$ 15.9 kg. This dataset will be used as training and testing data.

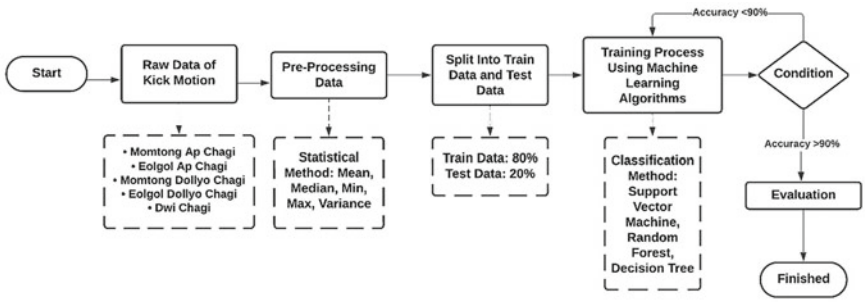


Fig. 1 Flowchart development model



Fig. 2 Taekwondo kick dataset retrieval process

2.2 Inertial Measurement Unit (IMU)

IMU, short for Inertial Measurement Unit, is an electronic device that combines several inertial sensors, such as an accelerometer, gyroscope, and magnetometer, into a single package. On the Arduino Nano BLE 33, there is an LSM9DS1 sensor which is included in this type of IMU sensor [9]. The LSM9DS1 can measure linear acceleration in three axes using the accelerometer, rotational angular velocity in three axes using a gyroscope, and magnetic field in three axes using a magnetometer. By combining data from the three sensors, the LSM9DS1 can provide detailed information about the movement and orientation of an object in three-dimensional space [10]. The LSM9DS1 sensor is generally used to develop motion recognition systems, measure position and orientation, and monitor human activity [11]. The data these sensors provide is very important in motion analysis and decision-making based on changes in objects' speed, acceleration and orientation [3, 8].

2.3 Preprocessing

Data reprocessing is done to convert raw data into data that is ready for use. The preprocessing stage in this study consisted of hot encoding, labeling, and cleaning. The hot encoding method is used to map the values of X, Y, and Z to numbers 0 and 1 and check for missing values [5, 10]. The labeling process is carried out to provide tags or labels to the data, especially for types of kicks. Before the labeling process is carried out, the dataset must go through the cleaning stage first. The cleaning process is done manually by deleting data patterns irrelevant to the kick movement and cleaning data such as a value of 0.

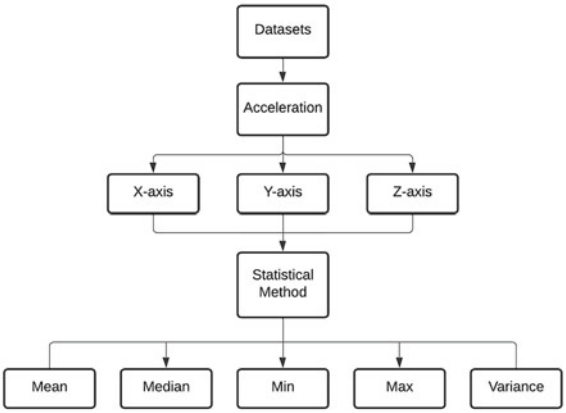


Fig. 3 Statistics features used

2.4 Feature Extraction Using Statistical Method

From the data of X, Y, and Z values that are owned, it will contain statistical processes that will assist in recognizing patterns in the form of mean, median, and others, as shown in Fig. 3. The feature extraction method using a statistical approach is one of the common approaches used in image processing and data analysis to obtain relevant features [12]. These statistical features are input in machine learning models for classification purposes [13]. This statistical feature extraction is useful for producing concise but informative data representations in the built model. The feature extraction process involves processing data from the X, Y, and Z axes into a data frame format.

In execution the three-axis values are used for processing because the movement technique that will be recognized in the form of a Taekwondo kick is a technique that is carried out paying attention to the direction of movement from above, below, right side, and left side. The windowing technique in data processing here refers to pattern recognition in every 130 data because one kick movement consists of 130 rows of numbers in the form of X, Y, and Z values.

This process is done by reading each frame using an open-source library. The iteration continues until all data is read. Some of the statistical features selected for use include the mean, median, min, max, and variance. These features help describe the characteristics of the data, the distribution of the data, the shape of the distribution, and so on. In this study, these statistical features play an important role in improving the predictions and performance of machine learning models [4, 29].

2.5 Hyperparameter Tuning with GridsearchCV

Parameter tuning technique, also known as hyperparameter optimization or tuning, is finding optimal hyperparameters to achieve the best performance. In this study,

the GridSearchCV method is used as a tuning technique to find the best parameters. GridSearchCV is commonly used to select optimal parameters for the Support Vector Machine (SVM) algorithm [15, 16, 28]. This method helps find the optimal combination of parameters for SVM by evaluating performance using cross-validation [17]. Before running the algorithm, the values for each property must be defined first. After determining the parameter values and doing the definition, the next step is fitting these parameters. In the Taekwondo kick dataset, 150 fittings are performed using threefold. The best parameters for the model were obtained through the SVM fitting process, namely C = 1000, gamma = 0.01, and kernel = RBF [18].

2.6 SVM Classifier

The implemented algorithm is Support Vector Machine (SVM) in this research. SVM is a technique to find hyperplanes that separate two datasets from different classes [13]. The hyperplane played an important role in separating the classes (Fig. 4). The successful implementation of this algorithm also depends on the parameters used [19]. This study's parameters include kernel type, C value, and gamma. Therefore, non-linear SVM is the most appropriate choice to overcome this classification problem [4, 14, 20, 26].

The types of kernels used in this study are as follows:

Linear. The linear kernel is the simplest SVM kernel. No nonlinear transformations are performed, so the linear kernel operates on data in the same dimensional space as the original data [19, 21, 29]. The formula of the linear kernel is as Eq. 1.

$$K(x, y) = x, y \tag{1}$$

x vector kernel value x
y vector kernel value y.

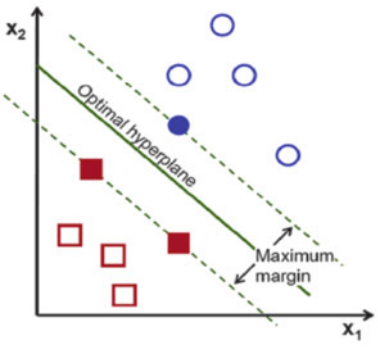


Fig. 4 Optimal hyperplane with maximum margin in SVM [13]

Radial Basis Function (RBF). The RBF kernel is the most commonly used kernel in SVM [22]. This kernel maps data into infinite dimensions using the Gaussian kernel function. The formula of the Gaussian kernel (RBF) is expressed as in Eq. 2.

$$K(xi, yi) = \exp\left(-\frac{\|xi - xj\|^2}{2\sigma^2}\right) \tag{2}$$

Xi vector in feature space x
Yj vector in feature space y
σ sigma, how sharp or how broad is the functionality of the RBF kernel.

2.7 MicroMLgen

MicroMLgen is an open-source framework designed to implement lightweight and efficient machine learning models on resource-constrained devices, such as microcontrollers and other edge devices. MicroMLgen's primary focus is optimizing model size and speed, as well as enabling real-time and local execution of machine learning without the need for an internet connection. Several machine learning methods supported by MicroMLgen include Decision Tree, Random Forest, XGBoost, GaussianNB, Support Vector Machine (SVM), Relevant Vector Machine, SEFR, and PCA. MicroMLgen also includes an optimizer that automatically optimizes machine learning models to reduce their size without sacrificing accuracy. This optimizer enables smaller machine learning models to run on edge devices by reducing memory usage and power requirements [7, 16].

2.8 Evaluation and Validation

Validation and evaluation are two crucial elements in Machine Learning. Validation is used to measure the ability of Machine Learning models to generalize patterns from data that have never been seen before. Validation helps assess whether the model can produce accurate and consistent predictions on new data. In this study, the Cross Validation method was used for model validation [23].

The evaluation aims to measure the performance and quality of the model that has been trained using relevant evaluation metrics. Common evaluation metrics used include accuracy, precision, recall, and F1-score. Accuracy indicates how well the model can make predictions correctly overall, while precision and recall provide more detailed information about the quality of predictions for certain classes. The F1-score is a metric that combines precision and recall to provide an overall rating of model quality. In this study, the confusion matrix is used as an evaluation method [17, 18].

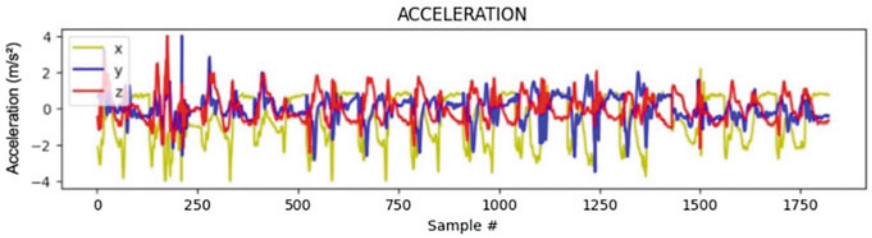


Fig. 5 Eolgol Ap Chagi dataset visualization right leg

3 Results

3.1 Dataset

The dataset taken is in the form of X, Y, and Z values from the kicking movements of Eolgol Ap Chagi, Momtong Ap Chagi, Eolgol Dollyo Chagi, Momtong Dollyo Chagi, and Dwi Chagi from the accelerometer sensor. There were 5 respondents who participated in the kick data retrieval. Respondents have different postures so that machine learning can learn a variety of data. The data taken was 70 kick movements, with 14 movements in each class consisting of the right and left foot. Retrieval of datasets is set with a threshold of 2.5 with the result that the data that has been cleaned reaches 130 lines per kick movement. The following is a display of one of the Taekwondo kick datasets [30] (Fig. 5).

3.2 Preprocessing

After going through the cleaning process, the data will be uploaded to Google Collab for analysis. Data that was initially separated will be integrated using the one hot encoding or label encoding method to combine into one integrated dataset. Next, the dataset will be divided into two parts, with 80% used as training data and 20% as testing data. After that, the Support Vector Machine (SVM) algorithm will be used to train the dataset [5].

3.3 Machine Learning Algorithm Selection

From the selection of the three methods previously described, the research will implement the SVM algorithm with an accuracy of 99.35% with five features. Table 1 displays the classification accuracy result.

Table 1 Accuracy results for right leg kick with selection features for training size 80%

Machine learning algorithm	Variance (%)	Median (%)	Mean (%)	Mean and variance (%)	Median and variance (%)	Mean and median (%)	Mean, median, min, max, variance (%)
Support vector machine	98	99	97	89	99	98	99
Random forest	96	82	84	70	88	88	94
Decision tree	79	99	96	89	99	98	99

3.4 Performance Measurement Using Support Vector Machine (SVM)

This experiment aims to test the SVM algorithm model in classifying five types of kicks in Taekwondo. The data used includes types of kicks such as Eolgol Ap Chagi, Momtong Ap Chagi, Eolgol Dollyo Chagi, Momtong Dollyo Chagi, and Dwi Chagi. Testing is carried out using pre-existing datasets. The model developed uses the RBF kernel type with parameters C = 1000 and gamma = 0.01. From this model, an accuracy of 99.35% was obtained. Table 2 displays the precision, recall, and f1-score result. While, Fig. 6 displays confusion matrix.

With an accuracy of 99.35%, there were still eleven prediction errors in the Dwi Chagi class, which was predicted Eolgol Dollyo Chagi. This can happen because of the similarity of the dataset patterns of the two types of kicks. If you look at the movement technique, these two movements are very similar. There is only a difference in the height of the feet, which, of course, changes the acceleration of the movement a little.

Table 2 Matrix evaluation

Kick Types	Precision (%)	Recall (%)	f1-score (%)
Momtong Ap Chagi	100	100	100
Eolgol Ap Chagi	100	100	100
Momtong Dollyo Chagi	100	100	100
Eolgol Dollyo Chagi	100	97	98
Dwi Chagi	97	100	98

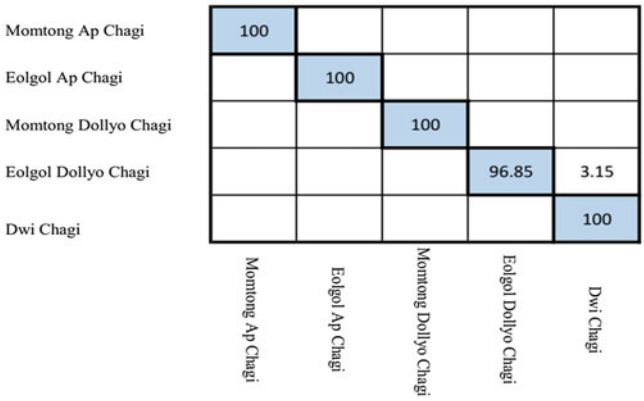


Fig. 6 Confusion Matrix (%)

3.5 Cross Validation

In this study, using CV = 3 for cross-validation, an average cross-validation score of 57% was obtained [24]. By dividing the data into three folds and training the model with those schemas, the average model accuracy is about 57%. This score describes how the model can make correct predictions on data that has never been seen before.

4 Discussion

Preprocessing using statistical methods proved to be effective in this study, especially considering that the machine learning model will be implemented on the Arduino nano 33 BLE development board.

In previous studies, results were also quite accurate, with 91.6% accuracy using the SVM method. However, the implementation development board is quite large and difficult for athletes. Therefore, it is replaced with Arduino Nano, which has small enough dimensions so that not all of the classification methods used can be used because it has limited use of the library in it.

The results of this study have significant implications for Taekwondo. Taekwondo has much equipment that must be used during matches or competitions. However, not many tools can automatically verify the type of kick to make it easier for the referee to judge. Furthermore, if this tool can continue to be developed and perfected in the future, it can already be used in matches without the role of a referee.

Although this research resulted in a high accuracy of 99.35%, it cannot be ignored that there were several cases of miss classification in the "Dwi Chagi" class, which can be seen in the confusion matrix Fig. 6. This can happen because the data pattern

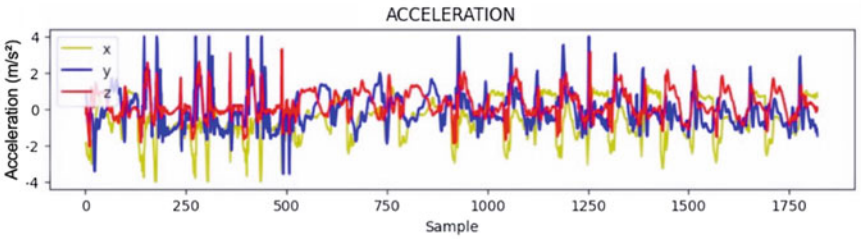


Fig. 7 Eolgol Dollyo Chagi dataset visualization right leg

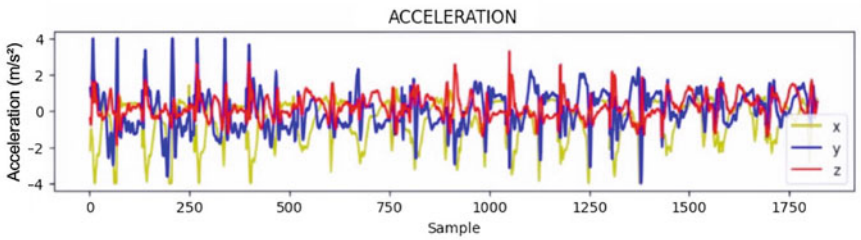


Fig. 8 Dwi Chagi dataset visualization right and left legs

between classes 0 (Momtong Ap Chagi) and 1 (Eolgol Ap Chagi) is similar. The following is a visualization of the two dataset classes (Fig. 7).

Even on the z-axis and x-axis you can see the similarities between the two datasets. Evaluation results with cross validation showed adequate results. With an accuracy rate of 57%, the model can generalize well to data that has never been seen before. It is possible to adjust the values of the C and gamma parameters to improve the model's performance. Even though hyperparameter tuning has been done using Gridsearch to find the best combination, there is still added value in trying different parameter variations to compare them and avoid overfitting (Fig. 8).

This study uses more class data and more advanced classifiers than previous studies [7]. Previous research used only two types of kicks, Ap chagi and Dolyo chagi, with KNN as the classifier. In addition, the device's dimensions are smaller and real-time data acquisition. The weakness of this study is that the initial determination of the movements to be recognized cannot be done automatically. In addition, the developed system has yet to determine kick speed; it only identifies the type of movement. Combining the IMU sensor and video processing can solve this speed measurement problem.

5 Conclusion

In this study, several machine learning methods were tested for kick classification in taekwondo using statistical characteristics. From the selection feature experiments, it has been tested that using five features effectively improves model performance for classification. The Taekwondo kick movement classification system produces an accuracy of 99.35% with the accelerometer sensor's X, Y, and Z axis values. The testing process results for data that is the same as the dataset produce good results, but the results are not good for the testing process outside of the dataset. This can be concluded due to the overfitting of the data model due to the value of C and gamma from GridSearch. Using more sophisticated classifier such as artificial neural network or deep learning and using two sensor on the right and left feet of the user with the results of the classification is sent directly to the mobile app and can be saved to view the history that has been carried out.

References

1. Hailong L (2021) Role of artificial intelligence algorithm for taekwondo teaching effect evaluation model. Journal of Intelligent & Fuzzy Systems. 40(2):3239–3250
2. Ferreira da Silva Santos J, Herrera-Valenzuela T, Franchini E (2019) Establishing frequency speed of kick test classificatory tables in male and female taekwondo athletes. Kinesiology 51(2):213–218
3. Hoang HTT, Ha CN, Nguyen DT, Nguyen TN, Huynh TN, Phan TT et al (2022) Poses classification in a Taekwondo lesson using skeleton data extracted from videos with shallow and deep learning architectures, pp 447–461
4. Asenov A (2020) Analysis of an online taekwondo competition, vol 92
5. Attal F, Mohammed S, Dedabrishvili M, Chamroukhi F, Oukhellou L, Amirat Y (2015) Physical human activity recognition using wearable sensors. Sensors (Switzerland). MDPI AG 15:31314–31338
6. Worsey MTO, Espinosa HG, Shepherd JB, Thiel DV (2019) Inertial sensors for performance analysis in combat sports: a systematic review. Sports. MDPI, vol 7
7. Dharmmesta RA, Jaya IGP, Rizal A, Istiqomah I (2022) Classification of foot kicks in Taekwondo Using SVM (support vector machine) and KNN (K-nearest neighbors) algorithms. In: 2022 IEEE international conference on industry 40, artificial intelligence, and communications technology (IAICT) [Internet]. IEEE, pp 36–41. https://ieeexplore.ieee.org/document/9887475/
8. Gede Pustika JI, Rainta Athallah D, Achmad R, Istiqomah I (2022) Application foot kick classification in Taekwondo with inertia sensor and machine learning. IEEE
9. Erfianto B, Rizal A (2022) IMU-based respiratory signal processing using cascade complementary filter method. Mailly F (ed) J Sens [Internet] 2022:1–16. https://www.hindawi.com/journals/js/2022/7987159/
10. eloquentarduino.github.com (2020) Introducing MicroML
11. Demrozi F, Pravadelli G, Bihorac A, Rashidi P (2020) Human activity recognition using inertial, physiological and environmental sensors: a comprehensive survey. IEEE Access. Institute of Electrical and Electronics Engineers Inc. 8:210816–210836
12. Ahmad GN, Fatima H, Ullah S, Salah Saidi A, Imdadullah (2022) Efficient medical diagnosis of human heart diseases using machine learning techniques with and without GridSearchCV. IEEE Access 10:80151–80173

13. Puspitasari A, Rizal A, Mukhtar H (2020) Prosthetic finger movement controller based on EMG signals using statistical feature and K-nearest neighbors. Int J Sci Technol Res 9(03):1472–1475
14. Fauzan MRA, Rizal A, Hadiyoso S (2022) ECG biometric using statistical feature of EEMD and VMD. In: 2022 IEEE international conference on industry 40, artificial intelligence, and communications technology (IAICT) [Internet]. IEEE, pp 113–118. https://ieeexplore.ieee.org/document/9887431/
15. Ismail I, Mukhtar H (2023) Development human activity recognition for the elderly using inertial sensor and statistical feature. In: Lecture notes in electrical engineering. Springer Science and Business Media Deutschland GmbH, pp 293–305
16. Pirjatullah KD, Nugrahadi DT, Muliadi FA (2021) Hyperparameter tuning using GridsearchCV on the comparison of the activation function of the ELM method to the classification of pneumonia in toddlers. In: 2021 4th international conference of computer and informatics engineering (IC2IE). IEEE pp 390–395
17. Anguita D, Ghelardoni L, Ghio A, Oneto L, Ridella S (2012) The 'K' in K-fold cross validation. In: ESANN 2012 proceedings, 20th European symposium on artificial neural networks, computational intelligence and machine learning
18. Mondal A, Saxena I, Tang H, Banerjee P (2018) A noise reduction technique based on nonlinear kernel function for heart sound analysis. IEEE J Biomed Health Inform [Internet] 22(3):775–784. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85046646037&doi=10.1109%2fJBHI.2017.2667685&partnerID=40&md5=b798ad7d6a5d50e61887343e0984b35c
19. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn [Internet] 20(3):273–297. http://link.springer.com/10.1007/BF00994018
20. Veeralagan J (2022) Hyper tuning using gridsearchcv on machine learning models for prognosticating dementia. https://doi.org/10.21203/rs.3.rs-2316713/v1
21. Zhang Y, Zhang X, Liu W, Luo Y, Yu E, Zou K et al (2014) Automatic sleep staging using multi-dimensional feature extraction and multi-kernel fuzzy support vector machine. J Healthc Eng [Internet] 5(4):505–520. http://www.ncbi.nlm.nih.gov/pubmed/25516130
22. Niu Y, Shang Y, Tian Y. Multi-view SVM classification with feature selection. Procedia Comput Sci [Internet] 162:405–412. https://doi.org/10.1016/j.procs.2019.12.004
23. Yadav S, Shukla S. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In: 2016 IEEE 6th international conference on advanced computing (IACC). IEEE, pp 78–83
24. Andersen T, Martinez T (1999) Cross validation and MLP architecture selection. In: IJCNN'99 international joint conference on neural networks proceedings [Internet]. IEEE, pp 1614–1619. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=832613
25. Tong L, Song Q, Ge Y, Liu M (2013) HMM-based human fall detection and prediction method using tri-axial accelerometer. IEEE Sens J 13(5):1849–1856
26. Styawati S, Mustofa K (2019) A support vector machine-firefly algorithm for movie opinion data classification. IJCCS (Indones J Comput Cybern Syst) 13(3):219
27. Ayumi V, Fanany MI, Fanany MI. A comparison of SVM and RVM for human action recognition multiple regularizations deep learning for paddy growth stages classification from LANDSAT-8 view project argumentation mining view project a comparison of SVM and RVM for human action recognition. https://www.researchgate.net/publication/281765294
28. Ranjan GSK, Kumar Verma A, Radhika S (2019) K-nearest neighbors and grid search CV based real time fault monitoring system for industries. In: 2019 IEEE 5th international conference for convergence in technology, I2CT 2019. Institute of Electrical and Electronics Engineers Inc.
29. Aditya Yanuar (2018) Statistical machine learning. www.mipa.ugm
30. Ishac K, Eager D (2021) Evaluating martial arts punching kinematics using a vision and inertial sensing system. Sensors 21(6):1–25
31. Lawi A, Aziz F, Wungo SL (2019) Increasing accuracy of classification physical activity based on smartphone using ensemble logistic regression with boosting method. J Phys Conf Ser. Institute of Physics Publishing

Classification EEG Signal Using Texture Analysis and Artificial Neural Network for Alcoholic Detection

Donny Setiawan Beu, Hilal Hamdi Simatupang, Achmad Rizal, Rita Purnamasari, and Yunendah Nur Fuadah

Abstract EEG is a powerful and popular technique for measuring brain activity, which reflects the condition of a person's brain. A person's brain health can be determined by monitoring brain activity used EEG technique. It has been demonstrated that EEG signals can be used as a diagnostic tool in evaluating individuals with alcoholism. Using the proper diagnostic method for EEG signals, the individual evaluating under alcoholism has been demonstrated. EEG signals record the brain's electrical activity, measured from the scalp. The measurements obtained from the EEG are used to confirm or rule out conditions such as alcoholism. Alcohol consumption is associated with specific patterns of brain electrical activity in adults, and the brain activity of individuals with alcoholism differs from non-alcoholics in several ways. This study proposes a feature extraction method for multichannel EEG signals using texture analysis for alcoholic and non-alcoholic classification. Multichannel EEG signal is treated as an image and processed using the texture analysis method. Then it is classified using an artificial neural network. The highest accuracy is achieved using the GLDM feature extraction method at a distance of 3 and an angle of 0°, resulting in an accuracy rate of 93.9%. The method used proved to be higher than previous studies using similar methods. The proposed method is expected to be used for other multichannel biomedical signal processing.

Keywords EEG · Classification · Alcoholic · Gry level co-occurrence matrix (GLCM) · Gray-level difference matrix (GLDM) · Artificial neural network (ANN)

D. S. Beu · H. H. Simatupang · A. Rizal (✉) · R. Purnamasari
School of Electrical Engineering, Telkom University, Bandung 40257, Indonesia
e-mail: achmadrizal@telkomuniversity.ac.id

Y. N. Fuadah
Computational Medicine Lab, Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_4

1 Introduction

Based on information in the Global Status Report on Alcohol and Health 2018, Indonesia's population reached 260,581,100 people [1]. The population noted that 0.8% had alcohol-related disorders, and 0.7% showed signs of alcohol dependence in both men and women [1, 2]. Alcoholism is a condition in which a person experiences alcohol dependence. Individuals who consume alcohol may struggle to control cravings and not pay much attention to the potential harm associated with such consumption [3]. Excessive alcohol consumption has very dangerous risks, including diseases such as cardiomyopathy, stroke, heart disorders, liver disorders, and an increased risk of cancer [4]. Long-term alcohol use can interfere with brain development. The short-term effects of alcohol can cause memory impairment, loss of consciousness, lack of alertness, and difficulty in decision-making [5]. Alcoholism not only negatively affects the individual who experiences it but can also have a negative impact on the surrounding environment due to behaviors influenced by alcohol [6].

Research on EEG signal processing for alcoholic detection has been done directly using the multichannel EEG signal processing method [7]. The development of Multichannel EEG signal processing methods continues to be carried out until now. Widadi using gamma signal of EEG for alcoholic detection in multichannel EEG [8]. Research by Ekaputri et al. [9] also recorded the highest accuracy results of 77.8%. In 2022, Aprilia and his colleagues researched using gray level difference matrix (GLDM) for Alcoholic EEG signals [10]. The study noted that logistic regression and linear discriminant analysis (LDA) methods using GLDM with an angle of 90° resulted in the highest accuracy value of 73.3% [10]. The classification process is carried out by processing EEG signals converted into images. These results can be compared with previous research by Barus et al., which achieved an accuracy of 75.25%, Recall of 78%, Precision of 73.93%, and F1 Score of 75.91% [11]. Texture analysis for new multichannel EEG signals results in low accuracy. Further research still needs to be done to explore alternative methods to improve accuracy.

Our research aims to improve the accuracy score on alcoholic and non-alcoholic image classification through the image enhancement process. The dataset used comes from the UCI Machine Learning Repository and consists of two classes that are the focus of the research [12]. Previously, similar studies have been conducted using the same dataset but have yet to involve improving image quality [10, 11]. Therefore, contributions of this study are:

1. Provide image processing method for multichannel EEG signal especially texture analysis.
2. We compare gray level difference matrix (GLDM) and gray level co-occurrence matrix (GLCM) as feature extraction method.
3. We use several distances of pixel for feature extraction, where in previous studies usually used $d = 1$.

In addition, we use artificial intelligent methods to obtain more optimal accuracy results. The final results of this research are expected to significantly contribute to image classification and its application to similar research.

2 Materials and Method

The signal processing in this study begins with setting the channel of the EEG signal so that it becomes an N channel x data sample as displayed in Fig. 1. Furthermore, the signal sample value is converted to be similar to the image, namely with a value of 0–255. Furthermore, the image quality is improved for further feature extraction using GLCM and GLDM. The next step is classification using an artificial neural network. Details of the steps are described in the next subsection.

2.1 Dataset

The data used can be obtained in UCI Machine Learning, where the data has been used by Widadi et al. [8]. The EEG data consists of 2 classes, namely alcoholic and normal or non-alcoholic. Each class contains 600 data captured from a frequency of 256 Hz for 1 s across 64 channels (Fig. 2). Sampling is done by providing stimulus 1 and stimulus 2.

There were 120 subjects, and each subject carried out 120 experiments. Demographic data from the subject (gender, age, height, weight, or any other health condition) is not available in the dataset used. The dataset obtained above was converted from one dimension to two dimensions by channel arrangement. The 64-channel EEG signal with 256 data samples will be treated as a 64 × 256 matrix image after the scaling process into 0–255 value as part of image conversion process [9].

2.2 Preprocessing

Preprocessing is focused on improving image quality. Image enhancement is an image preservation process that is easier for the human eye to interpret [13]. The method used in this paper is histogram equalization. Histogram equalization is image processing to increase the overall image contrast. In the equalization histogram technique, the distribution of image intensity values is changed to achieve an even balance [14].



Fig. 1 Diagram block of proposed method

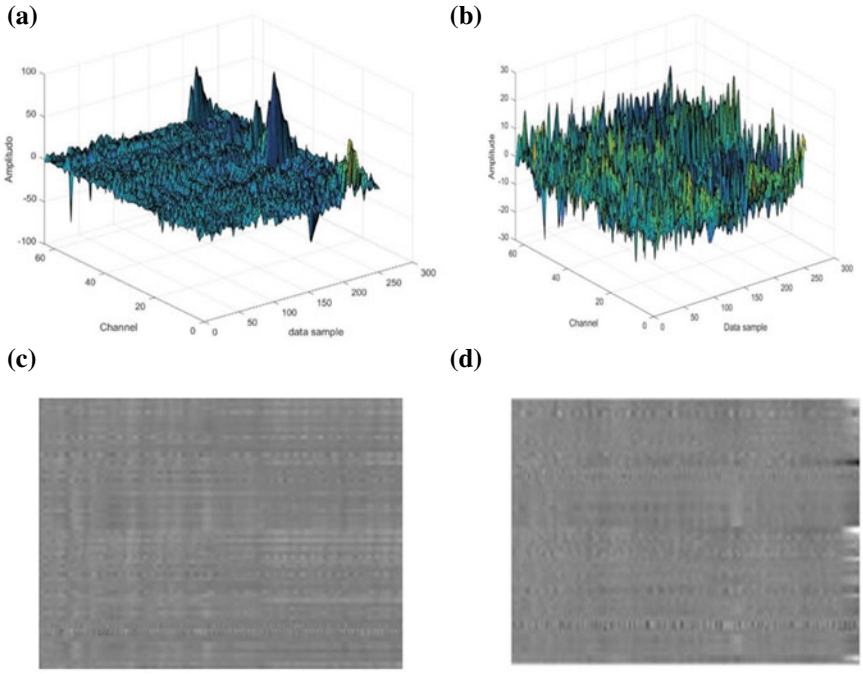


Fig. 2 **a** Plot of 64 channel EEG alcoholic signal. **b** Plot of 64 channel EEG non-alcoholic signal. **c** Image conversion of EEG alcoholic signal. **d** Image conversion of EEG non-alcoholic signal

2.3 Feature Extraction

Feature extraction is a feature vector interpreted as an n-dimensional vector encoding for indexing content and image databases [15]. Feature extraction can be used to take features from the object you want to study. In this study, the texture of the 64-channel EEG signal that has been converted to an image will be calculated [16]. The feature extraction process with texture analysis was carried out using two methods, namely the gray level co-occurrence matrix (GLCM) [17] and the gray level difference matrix (GLDM) [18].

Gray Level Co-occurrence Matrix (GLCM). Gray level co-occurrence matrix (GLCM) is used to evaluate texture or feature extraction [17]. GLCM determines the combined frequency of the specified pixel brightness values, thus representing the frequency formation of a pair of pixels [19]. The GLCM matrix shows the frequency of occurrence of a pair of two pixels with a certain intensity at a certain distance and direction in the image. If $h(i_1, i_2|\theta)$ is co-occurrence matrix for θ direction tha has the same information. While $p(i_1, i_2)$ is joint conditional probability density (JCPD) [20]. For calculating the feature of the image based on GLCM, we used several equations as presented in Eqs. 1–4 [18].

$$Contrast = \sum_{i_1} (i_1 - i_2)^2 p(i_1, i_2) \tag{1}$$

$$Homogeneity = \sum_{i_1} \sum_{i_1} \frac{p(i_1, i_2)}{1 + |i_1 - i_2|} \tag{2}$$

$$Energy = \sum_{i_1} \sum_{i_1} p^2(i_1, i_2) \tag{3}$$

$$Entropy = -\sum_{i_1} \sum_{i_1} p(i_1, i_2) \log p(i_1, i_2) \tag{4}$$

Co-occurrence pairs of pixels have an angular orientation of t and a distance of d. Pixels denote distances, and degrees denote angles. The angular orientation shape is based on the four angular directions, 0°,45°, 90°, and 135°, and 1 to 5-point distances between pixels [21].

Gray Level Difference Matrix (GLDM). The gray level difference matrix (GLDM) method calculates the absolute difference of a pair of pixels at a certain distance and direction [22]. The calculated directions are 0°, 45°, 90°, and 135°. Pixel difference is calculated horizontally, vertically, and diagonally and has different characteristics depending on the image's gray level. If $H(g|\theta)$ is a number of pixels that has g differences in θ direction, then we have $h(g|\theta) = H(g|\theta)/\sum H(g|\theta)$ as probability of adjacent pixel has difference of g when scanned in θ direction [20, 23]. We used several features generated from GLDM as expressed in Eqs. (5)–(9).

$$Gradient\ Contrast = GC = \sum_{g} g^2(g|\theta) \tag{5}$$

$$Gradient\ Second\ Moment = GSM = \sum_{g} [h(\theta)]^2 \tag{6}$$

$$Gradient\ Second\ Moment = GSM = \sum_{g} [h(\theta)]^2 \tag{7}$$

$$Gradient\ Mean = GE = h(g|\theta)g \tag{8}$$

$$Inverse\text{-}difference\ moment = IDM = \sum_{g} \frac{h(g|\theta)}{(g^2 + 1)} \tag{9}$$

As in GLCM, GLDM is also calculated at a distance of 1–5 pixels and direction $\theta = 0°, 45°, 90°$, and 135°.

2.4 *Classifier*

Classification is a technique used to categorize members into predetermined categories. These members are placed in certain categories based on the similarity of data characteristics [24]. Machine learning generally refers to a class of algorithms that can create predictive models automatically by recognizing patterns in data [25]. In this study, the classification method used is Artificial Neural Network (ANN). ANN or Artificial Neural Network, is a computational model inspired by the structure and function of human biological neural networks. ANN consists of processing units called neurons or nodes, interconnected through connections with adjustable weights [26]. One of the advantages of ANN is its ability to model linear and non-linear patterns with a wide range. In addition, ANN also can identify possible interactions between predictive variables [27]. The performance of the calcification is calculated using Eq. (10) where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

3 Results

Figure 3 displays the Histogram equalization (HE) results of the converted image from the N-channel EEG plot. Visually, it can be seen that the image's contrast has increased compared to before. While for testing classification performance, we tested using 80% training data and 20% test data. The accuracy result is shown in Table 1. This study tested four features in GLCM and five features in GLDM with four directions and distance d = 1–5. Another parameter tested is the epoch value of the ANN. The epoch values used in this test are 25, 50, and 100 [28]. We did not explore ANN parameters in this study. The results in Table 1 are the highest accuracy achieved by ANN.

In Table 1, the highest accuracy is 93.9% using GLDM at a distance of d = 3 and an angle of 0°. It can be seen that the selection of different distances produces



(a) **(b)**

Fig. 3 **a** Converted image before histogram equalization. **b** Converted image after histogram equalization

Table 1 Classification accuracy at various distances and directions for GLCM and GLDM features

Distances	Direction	Accuracy (%)	
		GLCM	GLDM
1	0°	68.1	88.9
	45°	74	89.2
	90°	75.4	81.6
	135°	76.3	85.8
2	0°	70.5	82.1
	45°	63.7	83.1
	90°	84.6	87.5
	135°	83.5	88.3
3	0°	79.5	93.9
	45°	83.7	88.5
	90°	83.8	85.1
	135°	80.4	88.5
4	0°	80.5	88.8
	45°	80.8	90.2
	90°	86.8	89.8
	135°	79.2	91.7
5	0°	81.5	89.8
	45°	79	84.6
	90°	91	81.6
	135°	81	83.3

different accuracy. There is no regular pattern regarding the accuracy of changes in angle and distance.

4 Discussion

GLCM and GLDM, which give different results in this test, indicate that the two methods have different approaches to extracting texture characteristics from images. In this test, the different results between GLCM and GLDM may be due to differences in how to capture and describe texture features in images. GLCM bases its analysis on the occurrence and spatial relationships between pixel pairs, whereas GLDM focuses on their absolute differences [29]. Another possibility is the influence of the direction of orientation and spatial distance used in GLDM [30]. Different direction and distance choices may produce different results in measuring pixel differences. In these tests, different results may be caused by the use of certain directions and distances that may be more suitable or more sensitive to the type of image being tested.

Classification using ANN gives high results. The highest accuracy is obtained using GLDM feature extraction at a distance of 3 directions 0° with an accuracy value of 93.9%. The 0° direction gives the best results because the direction follows the sequence of one channel. The research was carried out at a distance of 1–5, so the signal change looks very significant.

Compared to previous studies with the same approach and dataset [10, 11], this study was carried out more comprehensively because it uses image enhancement methods to improve the quality of converted images from 64 channels of EEG signals. In addition to feature extraction: GLCM and GLDM are performed at several pixel spacings, not just one spacing [10]. Previous research only used GLDM as a feature extraction method. The weakness of this study is that there needs to be parameter optimization in the ANN used. This leaves an opportunity to improve the performance of the proposed method by parameter optimization, feature selection, or using more advanced classifiers.

5 Conclusion

This research proposes a multichannel EEG feature extraction method for the classification of alcohol and non-alcoholic using the texture analysis method. The EEG signal in several N channels is converted into an image, and then an enhancement is carried out. For texture analysis, the GLCM and GLDM methods are used. Feature extraction is carried out in several directions and at several pixel distances. The highest accuracy of 93.9% was obtained from the test using GLDM with $d = 3$ and a direction of 0°. These results are better than previous studies using GLDM. The proposed method is expected to be an alternative for other biomedical signal processing, not only for EEG signals. Furthermore, further exploration can be carried out using different forms of EEG signal conversion or applying different feature extraction methods.

References

1. World Health Organization (2018) Global status report on alcohol and health 2018. Switzerland
2. Mukhtar H, Qaisar SM, Zaguia A (2021) Deep convolutional neural network regularization for alcoholism detection using EEG signals. Sensors 21(16):5456
3. Farsi L, Siuly S, Kabir E, Wang H (2021) Classification of alcoholic EEG signals using a deep learning method. IEEE Sens J 21(3):3552–3560
4. Li H, Wu L (2022) EEG classification of normal and alcoholic by deep learning. Brain Sci 12(6):778
5. Purnamasari PD, Ratna AAP, Kusumoputro B (2017) Relative wavelet bispectrum feature for alcoholic EEG signal classification using artificial neural network. In: QiR 2017—2017 15th international conference on quality in research (QiR): international symposium on electrical and computer engineering

6. Buriro AB, Ahmed B, Baloch G, Ahmed J, Shoorangiz R, Weddell SJ et al (2021) Classification of alcoholic EEG signals using wavelet scattering transform-based features. Comput Biol Med 139:104969
7. Saddam M, Tjandrasa H, Navastara DA (2018) Classification of alcoholic EEG using wavelet packet decomposition, principal component analysis, and combination of genetic algorithm and neural network. In: Proceedings of the 11th international conference on information and communication technology and system, ICTS 2017, 2018 January, pp 19–24
8. Widadi R, Soesanti I, Wahyunggoro O (2018) EEG classification using elliptic filter and multi-layer perceptron based on Gamma activity features. In: Proceedings - 2018 4th international conference on science and technology, ICST 2018. IEEE, pp 1–5. https://ieeexplore.ieee.org/document/8528568/
9. Ekaputri C, Widadi R, Rizal A (2020) EEG signal classification for alcoholic and non-alcoholic person using multilevel wavelet packet entropy and support vector machine. In: 2020 8th international conference on information and communication technology (ICoICT) [Internet]. IEEE, pp 1–4. https://ieeexplore.ieee.org/document/9166233/
10. Aprillia BS, Rizal A, Arik MG (2022) Gray level differences matrix (GLDM) for alcoholic EEG signal classification. In: The 4th 2021 international conference on applied information technology and innovation (ICAITI 2021)
11. Barus DT, Masri F, Rizal A (2020) NGBoost interpretation using LIME for alcoholic EEG signal based on GLDM feature extraction. Adv Intell Syst Comput 1294:894–904
12. Zhang XL, Begleiter H, Porjesz B, Wang W, Litke A (1995) Event related potentials during object recognition tasks. Brain Res Bull 38(6):531–538
13. Erwin E, Kiyatmoko T (2018) Retinal vessel extraction using dynamic threshold and enhancement image filter from retina fundus. J Inf Syst Telecommun 6(4):189–196
14. Wakaf Z, Jalab HA (2018) Defect detection based on extreme edge of defective region histogram. J King Saud Univ Comput Inf Sci [Internet] 30(1):33–40. https://doi.org/10.1016/j.jksuci.2016.11.001
15. Guyon I, Gunn S, Nikvares M, Zadeh L (2006) Feature extraction, foundations and applications. Zadeh L (ed) Physica-Verlag, Springer
16. Kairuddin WNHW, Mahmud WMHW (2017) Texture feature analysis for different resolution level of kidney ultrasound images. In: IOP conference series: materials science and engineering. Institute of Physics Publishing
17. Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. IEEE Trans Syst Man Cybern 3(6):610–621
18. Weszka JS, Dyer CR, Rosenfeld A (1976) A comparative study of texture measures for terrain classification. IEEE Trans Syst Man Cybern SMC-6(2):269–285
19. Fuadah YN, Setiawan AW, Mengko TLR (2015) Performing high accuracy of the system for cataract detection using statistical texture analysis and K-nearest neighbor. In: 2015 international seminar on intelligent technology and its applications performing, pp 85–88
20. Goshtasby A (2012) Image registration, principles, tools and methods, 1st ed. Springer, London Limited
21. Mohanty AK, Beberta S, Lenka SK (2011) Classifying benign and malignant mass using GLCM and GLRLM based texture features from Mammogram. Int J Eng Res Appl (IJERA) 1(3):687–693
22. Hadiyoso S, Rizal A (2021) Empirical mode decomposition and grey level difference for lung sound classification. Traitement du Signal [Internet] 38(1):175–179. http://www.iieta.org/journals/ts/paper/10.18280/ts.380118
23. Najarian K, Splinter R (2012) Biomedical signal and image processing, 2nd ed. CRC Press Book
24. Fei Y, Li W (2017) Improve artificial neural network for medical analysis, diagnosis and prediction. J Crit Care [Internet] 40:293. https://doi.org/10.1016/j.jcrc.2017.06.012
25. Bardou D, Zhang K, Ahmad SM (2018) Lung sounds classification using convolutional neural networks. Artif Intell Med [Internet] 88:58–69. https://doi.org/10.1016/j.artmed.2018.04.008

26. Sarić R, Jokić D, Beganović N, Pokvić LG, Badnjević A (2020) FPGA-based real-time epileptic seizure classification using artificial neural network. Biomed Signal Process Control 62:102106
27. Amato F, López A, Peña-Méndez EM, Vaňhara P, Hampl A, Havel J (2013) Artificial neural networks in medical diagnosis. J Appl Biomed 11(2):47–58
28. Abdul-wadood DN, George LE, Rasheed NA (2014) Texture classification using artificial neural network for diagnosis of skin cancer. Int J Electron Commun Comput Eng 1(2):20–25
29. Mamli S, Kalbkhani H (2019) Gray-level co-occurrence matrix of Fourier synchro-squeezed transform for epileptic seizure detection. Biocybern Biomed Eng [Internet] 39(1):87–99. https://doi.org/10.1016/j.bbe.2018.10.006
30. Rizal A, Hidayat R, Nugroho HA (2018) Modification of grey level difference matrix (GLDM) for lung sound classification. In: 2018 4th international conference on science and technology (ICST) [Internet]. IEEE, pp 1–5. https://ieeexplore.ieee.org/document/8528650/

CC1101 Network for Healthcare Cyber Physical System on Air Quality Data Acquisition



Firmansyah Maulana Sugiartana Nursuwars, Nurul Hiron, Aldy Putra Aldya, and Angga Setiawan Wahyudin

Abstract There is currently no dedicated network for IoT needs on the Siliwangi University campus, IoT device data communication still relies on the wifi network for internet network needs, which means that the quality of data transmission will be compromised if the network is overloaded with internet users. Siliwangi University needs a suitable, practical, and good network so that it can be used as a data communication line by the installed IoT devices. The viability of setting up an interconnection network on the Siliwangi University campus needs to be examined. In this study, the effectiveness of the CC1101 network used by Siliwangi University to monitor air quality via the Internet of Things was evaluated. The system consists of two components: nodes and gateways. The nodes use Arduino Nano microcontroller boards to transmit air quality data to the gateway, while the gateway uses a nodeMCU ESP8266 microcontroller board to transmit the data to the internet. According to the test findings for the CC1101 using the 433 MHz frequency, the most data that can be transferred is 64 bytes, and it can transmit data up to a distance of 68 m under line-of-sight conditions and 40 m under non-line-of-sight conditions. The data collected shows that there is an average data transmission delay of 859.698 ms and a packet loss of 3.12%.

Keywords Air quality · CC1101 · ESP8266 · 433 MHz · IoT

F. M. S. Nursuwars (✉)
Department of Informatics, Siliwangi Universitas Tasikmalaya, Tasikmalaya, Indonesia
e-mail: firmansyah@unsil.ac.id

N. Hiron · A. S. Wahyudin
Department of Electrical Engineering, Siliwangi Universitas Tasikmalaya, Tasikmalaya, Indonesia

A. P. Aldya
Department of Information System, Siliwangi Universitas Tasikmalaya, Tasikmalaya, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_5

1 Introduction

Siliwangi University Tasikmalaya is one of the largest universities in East Priangan and is a state university in the city of Tasikmalaya [1, 2]. To monitor the condition of the campus environment, Siliwangi University requires an appropriate, feasible, and good network so that it can be used as a data communication line by IoT devices located at Siliwangi University, currently on the Siliwangi University campus there is no independent network available for IoT needs, so it is necessary to analyze and implement the interconnection network and what devices are suitable for use on the Siliwangi University campus such as CC1101.

CC1101 is one of the wireless data communication technologies, which utilizes low frequencies that can cover a fairly long communication distance and minimal power consumption [3]. CC1101 belongs to the IEEE 802.15 family of standards along with Bluetooth (802.15.1) and WPAN (802.15.3) with standard code IEEE 802.15.4 [4].

The use of wireless communication systems as a medium of communication on computer network systems or devices is increasingly popular today. Applications of wireless technology make information and communication fast and easy [5]. The technology created is a temperature and humidity monitoring system at several points in different areas. This system supports sensors at several points, the acquisition results will be sent to the internet wirelessly.

Research related to this article has been conducted on the problem of how to improve the accuracy of temperature sensors in IoT-based systems with the linear regression method [6]. Furthermore, research that discusses the implementation of ESP8266 in smart home systems [7]. And the utilization of the NarrowBand IoT (NB-IoT) module as a data communication module in the IoT-based rice weighing system [8]. The research has not explained the data communication module that is suitable for indoor communication and is different from the general communication used for internet network hotspots.

Based on these problems, this study analyzes the quality of the CC1101 network on an IoT-based air quality monitoring system at Siliwangi University. This study aims to study and apply the internet of things using the CC1101 module and the MQTT protocol at Siliwangi University.

2 Materials and Method

2.1 Materials

This section describes the devices and sensors used for the data collection process. The materials used are verified for feasibility by measuring data retrieval errors on sensors, device function performance by conducting unit tests and network feasibility by measuring delay, packet loss and RSSI values on wireless devices.

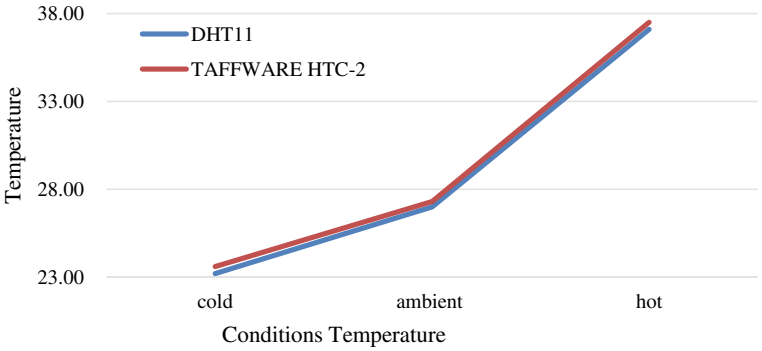


Fig. 1 Accuration temperature sensor

CC1101. The CC1101 is a low-cost sub-1 GHz transceiver designed for ultra-low power wireless applications. The circuit is primarily intended for the ISM (Industrial, Scientific and Medical) and SRD (Short Range Device) frequency bands at 315, 433, 868, and 915 MHz, but can be easily programmed for operation at other frequencies at 300–348, 387–464, and 779–928 MHz [3].

DHT11. The DHT11 sensor is a temperature and humidity sensor where the output of this sensor is a digital signal. The resolution of the DHT11 for temperature is 8 bits, minimum accuracy ±10 °C and maximum accuracy ±2 °C with a temperature measurement range from 0 to 50 °C [9].

Testing the DHT11 temperature sensor aims to determine the accuracy of the data acquisition process. In testing the DHT11 sensor is tested with 3 conditions, namely room temperature conditions, hot conditions and cold conditions. The comparison tool in testing the temperature and humidity sensor (DHT11) is using the Taffware tool (HTC-2). With an error rate of 1.69% in cold temperature conditions, 1.10% in ambient conditions and 1.07% in hot conditions, show in Fig. 1. With an average error rate of 1.29% so that it can be declared accurate for each measurement condition.

2.2 Research Methodology

The research methodology in this study is shown in Fig. 2. Research methodology consists of 7 stages; Literature study; System design; Material collection; System build; System testing; Data collection; Evaluation and Results.

Literature study is carried out to find references from journals, books, and other media in obtaining theories about research problems, research problems become the basis for system design, then collect materials as needed materials collected in the raft according to the design. The system design is in accordance with the system architecture in Fig. 3. The system consists of 2 nodes, the first node as a gateway node as shown in Fig. 4a and the second node as a client node as shown in Fig. 4b.

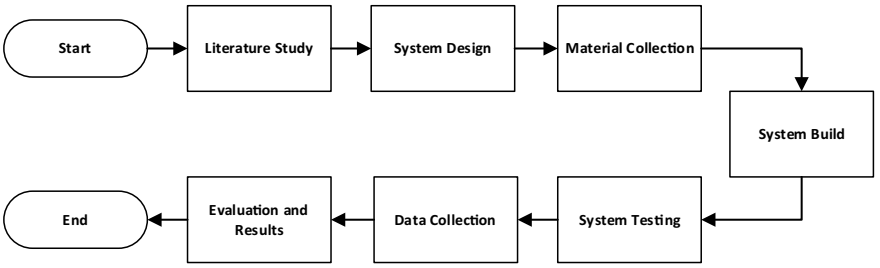


Fig. 2 Research methodology

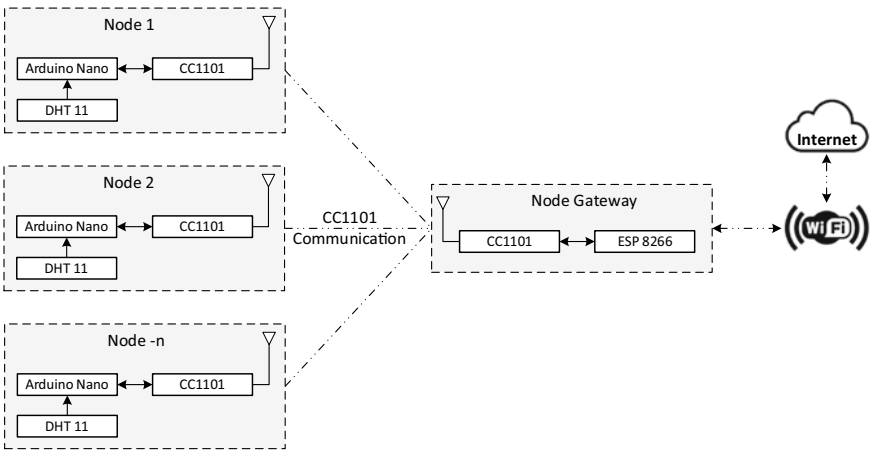


Fig. 3 System architecture

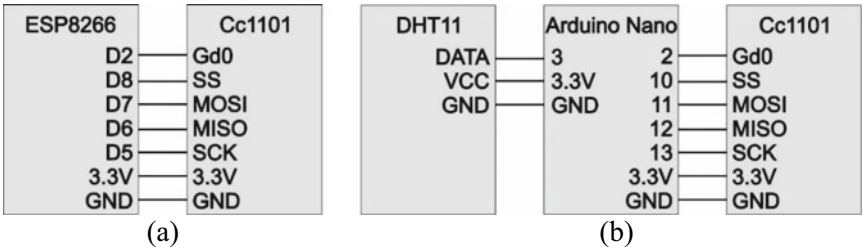


Fig. 4 Wiring diagram **a** gateway node, **b** client node

After the system is working, testing is carried out the data taken in the form of the ability how many bytes of data can be sent, how far the signal can be captured in line of signt and no line of signt conditions, as well as the presence of interference from other signals. The data obtained are evaluated and then conclusions are made to obtain results.

Delay. Delay is the time required from sending the sent data packet until the sent data packet [10–12]. Delay calculations can be formulated with Eqs. (1) and (2).

$$Delay = \frac{(Pt - Pk)}{Pk} \tag{1}$$

$$Delay\ Averge = \frac{\sum Delay}{\sum Pk} \tag{2}$$

where:

Pt: Package Sent
Pk: Package received
$\sum Delay$: Total *delay*
$\sum Pk$: Total packets received.

Packet Loss. Packet Loss is a test parameter that describes the condition of the number of packets lost during the transmission process from transmitter to receiver [13, 14]. The packet loss calculation can be formulated with Eq. 3.

$$Packet\ Loss = \frac{\sum Pk - \sum Pt}{\sum Pk} \times 100\% \tag{3}$$

where:

$\sum Pk$: Number of packages received
$\sum Pt$: Number of packages sent.

RSSI. RSSI is a unitless metric measurement to measure the signal strength received by radio or wireless devices [15–17]. RSSI is an indicator of the power value received by a wireless device at the receiver from the signal emitted by the transmitter device. By knowing the RSSI value we will know the quality of service of a wireless system in a particular place.

3 Result

3.1 Data Packet Testing

Testing the CC1101 unit to determine the performance of the function of how much data packets can be sent by the device from the node to the gateway [2, 18]. Testing is done by sending data starting from 1 byte until the data cannot be sent. Tests were carried out on the CC1101 unit which was set to Frequency 433 MHz, Frequency 868 MHz and Frequency 915 MHz with results such as Table 1, where the CC1101 module can send data packets up to 63 Bytes [19].

Table 1 CC1101 performance send data packet

No	Sent data width (Byte)	Received data width (Bytes)	Result (%)
1	1	1	Success
2	7	7	Success
3	14	14	Success
4	21	21	Success
5	28	28	Success
6	35	35	Success
7	42	42	Success
8	49	49	Success
9	54	54	Success
10	63	63	Success
11	64	0	Failed

3.2 CC1101 Distance Testing

This test was carried out in several ways, namely: with LOS (line of sight) conditions, not LOS, interference channel with LOS conditions, interference channel with non-LOS conditions. This test was also carried out with several frequencies, namely 433, 868, and 915 MHz.

LOS condition. LOS communication refers to the situation where two communication devices, such as antennas or wireless devices, have an unobstructed line of sight to each other [20–22]. In LOS conditions, radio signals or electromagnetic waves can travel from one device to another without physical barriers between them Tests under LOS (Line of Sight) conditions are illustrated in Fig. 5 and the test results are presented in Table 2.

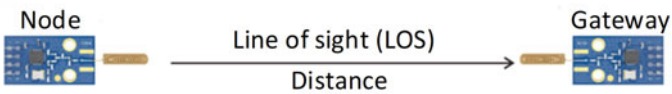


Fig. 5 Illustration of LOS (line of sight) conditions

Table 2 CC1101 distance LOS (line of sight) condition

No	Frequency (MHz)	Distance (m)
1	433	70
2	868	20
3	915	10

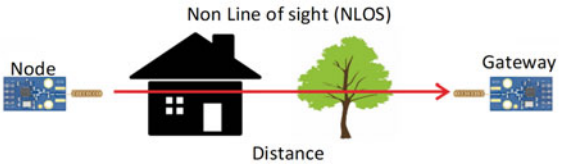


Fig. 6 Illustration of non line of sight (NLOS) condition

Table 3 CC1101 distance non line of sight (NLOS) condition

No	Frequency (MHz)	Distance (m)
1	433	40
2	868	10
3	915	10

Condition Non-LOS. Non-line of sight (NLOS) communication is the condition where there are physical obstacles or other hindrances that block the direct line of sight between two communication devices [20, 21, 23–26]. These barriers can be buildings, hills, trees, or other objects that obstruct signal transmission. Tests with non-LOS conditions are illustrated in Fig. 6 and the test results are presented in Table 3.

3.3 Channel Interference

Channel Interference LOS Condition. The interference channel test under LOS conditions with one additional gateway with the same channel [27] is illustrated in Fig. 7 and the test results are presented in Table 4.

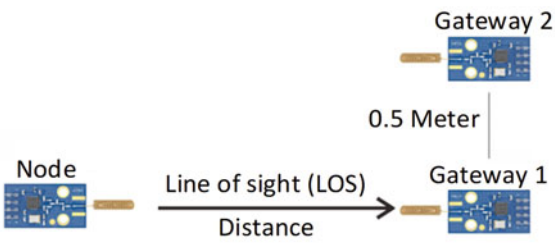


Fig. 7 Illustration of interference channel in LOS (line of sight) condition

Table 4 CC1101 distance LOS (line of sight) condition with channel interference

No	Frequency (MHz)	Distance (m)
1	433	45
2	868	20
3	915	10

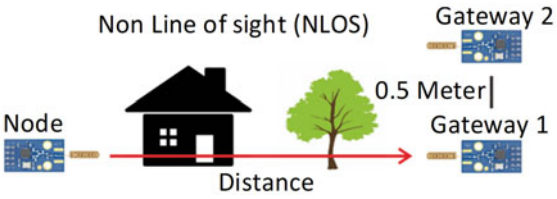


Fig. 8 Illustration of interference channel in non line of sight (NLOS) condition

Table 5 CC1101 distance non line of sight (NLOS) condition with channel interference

No	Frequency (MHz)	Distance (m)
1	433	30
2	868	10
3	915	5

Interference Channel Non-LOS Condition. The interference channel test under non-LOS conditions with one additional gateway with the same channel is illustrated in Fig. 8 and the test results are presented in Table 5.

3.4 CC1101 RSSI Quality Testing

This network quality test was carried out at Siliwangi University with the topology plan described in Fig. 9. This test was also carried out with several frequencies, namely: 433, 868 and 915 MHz. The data taken in this test are packet loss and average delay.

After collecting data according to the topology plan in Fig. 9, then Eq. 3 is used to obtain the total packet loss presented in Table 6. Table 6 shows the average packet loss is more than 90% for CC1101 data communications at the 868 and 915 MHz frequencies, which indicates more data was not received by the receiver when it was implemented at Siliwangi University, whereas the CC1101 network with a frequency of 433 MHz got a packet lost value of 3.28% so it could be declared suitable for implementation.

Test results the average delay of CC1101 by using Eqs. 2 and 3. The test results are presented in Table 7. Table 7 shows the test results of the average delay on the CC1101 network with a frequency of 433 MHz of 859.69 ms, while for the frequencies 868 and 915 MHz the value is infinite because the packet loss test results on these frequencies received a value greater than 90%. So it is stated that the CC1101 network is only suitable for implementation at Siliwangi University for the 433 MHz frequency only.

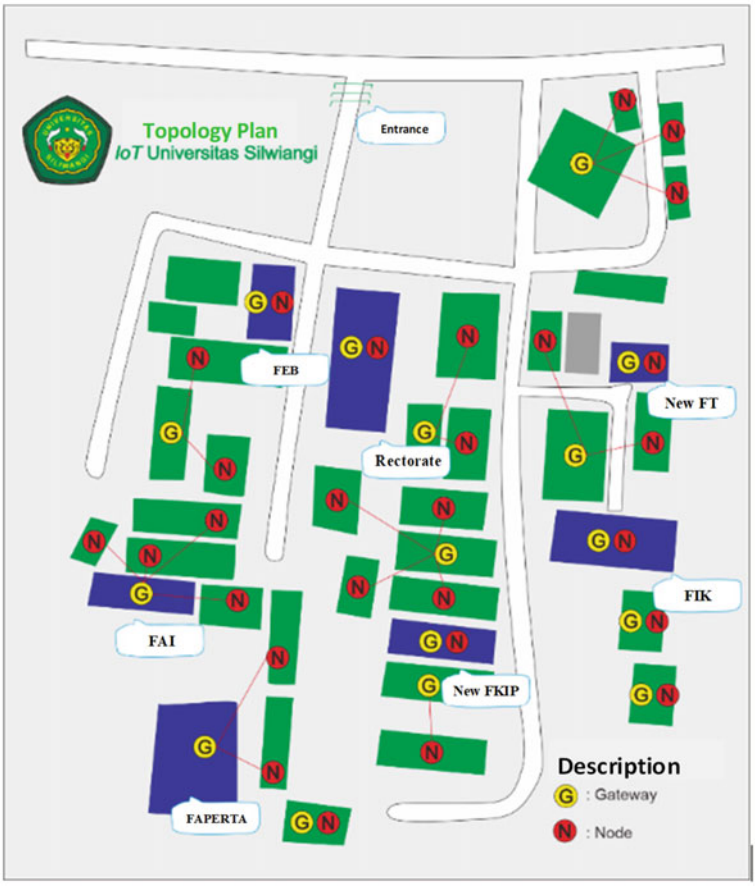


Fig. 9 CC1101 testing topology plan

4 Discussion

Based on the results of data collection monitoring of room temperature conditions with CC1101 according to Siliwangi University with the topology plan described in Fig. 11. there is still a very high packet loss in some areas when using frequencies of 868 and 915 MHz where according to Table 8 the packet loss reaches 100% with an average of 91.56% at frequencies of 868 MHz and 96.58% at frequencies of 915 MHz. it can be interpreted that the use of CC1101 with frequencies of 868 and 915 MHz is not feasible to implement at Siliwangi University. From these results, the next research process only uses CC1101 with frequencies of 433 MHz.

The quality of the 433 MHz frequency CC1101 network implemented at Siliwangi University for data communication networks in the air quality acquisition process is

Table 6 Packet loss testing result

No	Region name	Packet loss (%)		
		433 MHz	868 MHz	915 MHz
1	Rectorate building	2.00	100.00	100.00
2	FEB	5.00	100.00	100.00
3	FAI	3.00	100.00	100.00
4	FAPERTA	5.00	100.00	100.00
5	New FKIP building	8.00	100.00	100.00
6	Old FKIP building	5.00	100.00	100.00
7	Library	0.00	48.33	78.50
8	FIK	1.70	42.00	77.00
9	New FT building	6.00	100.00	100.00
10	Old FT building	7.00	100.00	100.00
11	Sports hall Mashudi	0.00	100.00	100.00
12	Electrical engineering laboratory	0.00	100.00	100.00
13	FISIP	0.00	100.00	100.00
Average		**3.28**	**91.56**	**96.58**

Table 7 Average delay testing results

No	Region name	Average delay (ms)		
		433 MHz	868 MHz	915 MHz
1	Rectorate building	886.27	∞	∞
2	FEB	864.05	∞	∞
3	FAI	851.33	∞	∞
4	FAPERTA	868.89	∞	∞
5	New FKIP building	864.36	∞	∞
6	Old FKIP building	860.87	∞	∞
7	Library	856.06	872.88	876.20
8	FIK	854.48	854.65	860.42
9	New FT building	853.29	∞	∞
10	Old FT building	860.94	∞	∞
11	Sports hall Mashudi	843.92	∞	∞
12	Electrical engineering laboratory	854.00	∞	∞
13	FISIP	857.62	∞	∞
Average		**859.69**	∞	∞

Table 8 RSSI testing results

No	Region name	Db
1	Rectorate building	−60
2	FEB	−64
3	FAI	−67
4	FAPERTA	−55
5	New FKIP building	−69
6	Old FKIP building	−66
7	Library	−50
8	FIK	−68
9	New FT building	−69
10	Old FT building	−64
11	Gor Mashudi	−53
12	Electrical engineering laboratory	−49
13	FISIP	−52

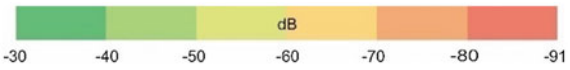


Fig. 10 CC1101 dB qualification color bar

described using color codes to represent RSSI quality based on dB values according to the qualification color bar in Fig. 10.

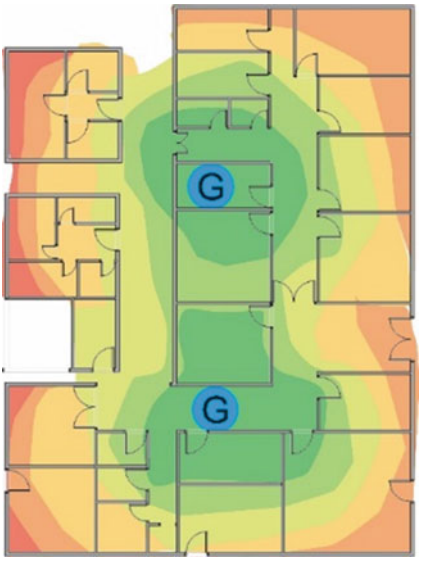


Fig. 11 Heatmap of the old faculty of engineering building

Figure 11 is also one of the heatmap results at Siliwangi University. As an example is the old faculty building, for the blue G icon is the gateway device. The results of network analysis based on RSSI taken when collecting data at Siliwangi University have the results of the RSSI graph at Siliwangi University. The results of the RSSI Value are presented in Table 8.

After analyzing the heatmap in each region, the results of packet loss and delay are obtained. Figures 12 and 13 present the results of packet loss and delay generated by CC1101.

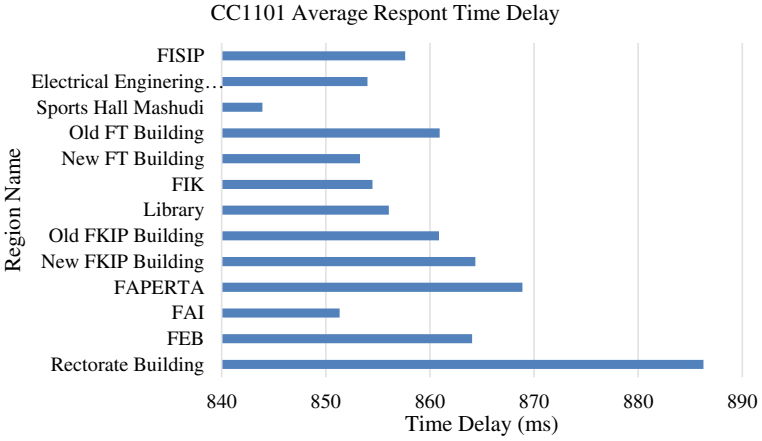


Fig. 12 Average delay graph of signal quality analysis results

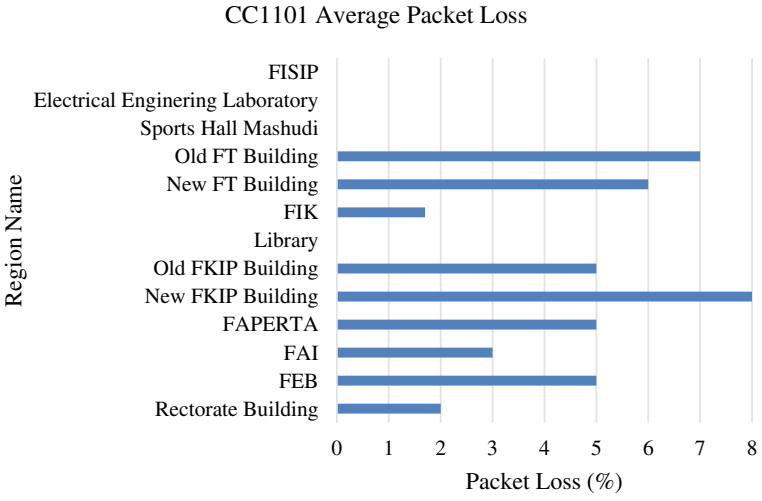


Fig. 13 Packet loss graph of signal quality analysis results

Results of testing analysis of packet loss, delay and how much data can be sent from the transmitting device to the receiving device. It can be determined that the CC1101 433 MHz network is suitable for implementation at Siliwangi University. The drawback of the CC1101 module is that data can only be sent in 63 bytes. If the data acquisition at various nodes exceeds 63 bytes, the data sending process must be split into several data packets so it will reduce the response time of IoT devices.

5 Conclusion

Based on the results of research and discussion, it can be concluded that the analysis of the quality of the CC1101 network on the IoT-based room temperature condition monitoring system at Siliwangi university, the results of the CC1101 testing maximum data that can be sent is 64 byte and can send data up to 68 m with line of sight conditions and 40 m with non-line of sight conditions using a frequency of 433 MHz. At a frequency of 868 MHz, it can send 63 bytes and is able to send data up to 20 m with line of sight conditions and 10 m with no line of sight conditions. At a frequency of 915 MHz, 63 bytes can be sent and able to send data up to 10 m with line of sight conditions and 10 m with no line of sight conditions. Based on data collection, the average delay obtained is 859.698 ms and packet loss is 3.12% at a frequency of 433 MHz. The success of CC1101 using the 433 MHz frequency for monitoring room temperature is 95%. So that CC1101 is suitable for use at Siliwangi University.

Frequency according to data collection and based on CC1101 testing, at a free frequency of 433 MHz, the average delay is 859.698 ms, while for a frequency of 868 MHz is infinity (∞), and for a frequency of 915 MHz is infinity (∞) the delay with an infinity value is because the data is not received by the client node. For packet loss at 433 MHz frequency is 3.23%, while for 868 MHz frequency is 91.56%, and 915 MHz frequency is 96.58%. So that the suitable frequency applied at Siliwangi University using CC1101 is 433 MHz.

References

1. Hidayat EW, Nursuwars FMS, Rahmatulloh Al (2017) Analisis service quality untuk mengukur tingkat kepuasan pengguna layanan tik di Universitas Siliwangi. In: Seminar Nasional Teknologi Informasi dan Multimedia. https://ojs.amikom.ac.id/index.php/semnasteknomedia/article/viewFile/1591/1583. Accessed 07 November 2022
2. Nursuwars FMS (2018) API (application programing interface) Mikrotik untuk Otentifikasi Sistem Akademik Universitas Siliwangi. Jurnal Siliwangi Seri Sains dan Teknologi 4(2)
3. Low-Power Sub-1 GHz RF Transceiver and E "cc1101." https://www.ti.com/lit/ds/symlink/cc1101.pdf. Accessed 07 November 2022
4. Hiron N et al (2016) Batch processing method in machine to machine wireless communication as smart and intelligent system. Int J Future Comput Commun 5(3):163–166. https://doi.org/10.18178/ijfcc.2016.5.3.464

5. Af'idah DI, Rochim AF, Widianto ED (2014) Perancangan Jaringan Sensor Nirkabel (JSN) untuk Memantau Suhu dan Kelembaban Menggunakan nRF24L01+. Jurnal Teknologi dan Sistem Komputer 2(4):267. https://doi.org/10.14710/jtsiskom.2.4.2014.267-276
6. Hariyanto T, Rahayu M, Satria F, Fadhlan MY (2019) Improving temperature sensor accuracy in the IoT trainer kit by linear regression method. In: Proceedings of the 2019 international conference on mechatronics, robotics and systems engineering, MoRSE 2019. https://doi.org/10.1109/MoRSE48060.2019.8998639
7. Fajar Wicaksono M (2017) Implementasi Modul Wifi Nodemcu Esp8266 Untuk Smart Home. Jurnal Teknik Komputer Unikom-Komputika 6(1)
8. Hayat MEAN, Nursuwars FMS, Aripin A (2022) Timbangan beras digital berbasis narrowband internet of things. J Energy Electr Eng (JEEE) 4(1)
9. Saptadi AH (2015) Perbandingan Akurasi Pengukuran Suhu dan Kelembaban Antara sensor DHT11 dan DHT22 Studi Komparatif pada platform ATMEL AVR dan Arduino. Jurnal Informatika, Telekomunikasi dan Elektronika 6(2). https://doi.org/10.20895/infotel.v6i2.73
10. Althoubi A, Alshahrani R, Peyravi H (2021) Delay analysis in IoT sensor networks†. Sensors 21(11). https://doi.org/10.3390/s21113876
11. Aljubayri M, Peng T, Shikh-Bahaei M (2021) Reduce delay of multipath TCP in IoT networks. Wirel Netw 27(6). https://doi.org/10.1007/s11276-021-02701-3
12. Sankayya M, Kumar Sakthivel R, Gayathri N, Al-Turjman F (2023) Wireless sensor network–based delay minimization framework for IoT applications. Pers Ubiquitous Comput 27(3). https://doi.org/10.1007/s00779-020-01517-w
13. Itahara S, Nishio T, Koda Y, Yamamoto K (2022) Communication-oriented model fine-tuning for packet-loss resilient distributed inference under highly lossy IoT networks. IEEE Access 10. https://doi.org/10.1109/ACCESS.2022.3149336
14. Chen YT, Lu LY, Yu X, Li X (2019) Adaptive method for packet loss types in IoT: an Naive Bayes distinguisher. Electronics (Switzerland) 8(2). https://doi.org/10.3390/electronics8020134
15. Zhu J, Hou P, Nagayama K, Hou Y, Denno S, Ferdian R (2022) Two-dimensional RSSI-based indoor localization using multiple leaky coaxial cables with a probabilistic neural network. IEEE Access 10. https://doi.org/10.1109/ACCESS.2022.3153083
16. Li G, Geng E, Ye Z, Xu Y, Lin J, Pang Y (2018) Indoor positioning algorithm based on the improved RSSI distance model. Sensors (Switzerland) 18(9). https://doi.org/10.3390/s18092820
17. Poulose A, Han DS (2021) Hybrid deep learning model based indoor positioning using wi-fi RSSI heat maps for autonomous applications. Electronics (Switzerland) 10(1). https://doi.org/10.3390/electronics10010002
18. Demircioğlu ED, Kalipsiz O (2022) API message-driven regression testing framework. Electronics (Switzerland) 11(17). https://doi.org/10.3390/electronics11172671
19. Stoitsov G (2014) One implementation of API interface for RouterOS. 3(2). https://doi.org/10.2772/42641
20. Fang C, Li S, Wang K (2022) Accurate underwater optical wireless communication model with both line-of-sight and non-line-of-sight channels. IEEE Photon J 14(6). https://doi.org/10.1109/JPHOT.2022.3216599
21. Alsaify BA, Almazari MM, Alazrai R, Daoud MI (2020) A dataset for Wi-Fi-based human activity recognition in line-of-sight and non-line-of-sight indoor environments. Data Brief 33. https://doi.org/10.1016/j.dib.2020.106534
22. Buller PF, McEvoy GM (2012) Strategy, human resource management and performance: sharpening line of sight. Human Resour Manag Rev 22(1). https://doi.org/10.1016/j.hrmr.2011.11.002
23. Heide F, O'Toole M, Zang K, Lindell DB, Diamond S, Wetzstein G (2019) Non-line-of-sight imaging with partial occluders and surface normal. ACM Trans Graph 38(3). https://doi.org/10.1145/3269977
24. Pan Z et al. (2022) Onsite non-line-of-sight imaging via online calibration. IEEE Photon J 14(5). https://doi.org/10.1109/JPHOT.2022.3207785

25. Merdan FAB, Thiagarajah SP, Dambul K (2022) Non-line of sight visible light communications: a technical and application based survey. Optik (Stuttg) 259. https://doi.org/10.1016/j.ijleo.2022.168982
26. Sun X et al (2020) Non-line-of-sight methodology for high-speed wireless optical communication in highly turbid water. Opt Commun 461. https://doi.org/10.1016/j.optcom.2020.125264
27. Ohaneme CO (2012) Analysis of interference and channel capacity in a CDMA wireless network using dynamic channel assignment (DCA) strategy. Int J Comput Netw Commun 4(5). https://doi.org/10.5121/ijcnc.2012.4510

Modelling of Solar Irradiance for Optimal Solar-Powered Car Performance at EPIC Solar Farm Pathway



Afidatul Nadia Mok Hat, Ruzlaini Ghoni, Mohd Tarmizi Ibrahim, Ahmad Firdaus Zali, and Fuaad Mohamed Nawawi

Abstract Solar irradiance is paramount when assessing solar panels' efficiency and overall effectiveness. The intensity and consistency of solar irradiance directly impact the performance and output of solar panels. The study aims to address the problem of understanding the impact of solar irradiance on the efficiency and performance of solar-powered cars. Specifically, the relationship between solar irradiance levels and the voltage output of solar panels installed on a car, focusing on the EPIC Solar Farm pathway in Kemaman, Terengganu. This research involved modeling solar irradiance and its impact on the performance of a solar-powered car. By analyzing the data measured solar irradiance, the research could correlate it with the voltage output of the solar panels installed on the car. The findings revealed a clear relationship between solar irradiance levels and the performance of the solar-powered car. The study highlighted the significance of accurate solar irradiance predictions for optimizing the performance of solar-powered cars. Car manufacturers and designers can enhance solar-powered vehicles' overall efficiency and range by considering solar irradiance levels at specific locations, such as the EPIC Solar Farm pathway. These results contribute valuable insights into the field of solar-powered transportation, showcasing the importance of solar irradiance modeling in developing more effective and sustainable solar-powered vehicles.

Keywords Solar irradiance · Solar-powered car · Renewable energy

A. N. M. Hat (✉) · R. Ghoni · M. T. Ibrahim
Faculty of Engineering Technology, University College TATI, Chukai, Malaysia
e-mail: afidatulnadia@uctati.edu.my

A. F. Zali · F. M. Nawawi
EPIC Solar Farm Sdn. Bhd., Kemaman Supply Base, Chukai, Malaysia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_6

1 Introduction

Solar energy is a clean, reliable, renewable source that can power vehicles [1]. Solar cells, which convert solar energy into electrical energy, are a promising technology for the future of sustainable energy. Solar-powered vehicles have been designed and tested, demonstrating their ability to start, and run using solar energy. The efficiency of solar cells plays a significant role in the range and driving capabilities of solar-powered vehicles. The impact of solar irradiance on solar-powered vehicles, such as cars, is crucial for optimizing their performance [2–5]. Highly efficient photovoltaic (PV) modules can enable solar-powered vehicles to drive significant distances without charging [6–12]. Therefore, developing high-efficiency and low-cost solar cells and modules is essential for the automotive application of solar energy. As solar energy is highly dependent on the intensity and duration of sunlight, it is essential to understand the effects of various weather conditions on solar-powered cars' energy output and efficiency.

Previous research studied the impact of solar irradiance on the performance of solar-powered cars by proving the higher solar irradiance is on the car roof [13], but this result depends on the car's route through which is open air, urban, or high-rise section [14]. On the contrary, a feasibility study conducted on the performance of solar-powered electric cars in four countries from the pattern of monthly solar irradiation showed the most effective performance of solar-powered electric cars solely by photovoltaic is the country that has the higher solar irradiance the whole year compared to the others [15, 16].

In summary, solar irradiance significantly impacts the energy efficiency and performance of solar-powered cars. Existing studies have often generalized findings without considering the specific solar irradiance conditions at the location. Therefore, it is essential to study the solar irradiance in the specific location of the operation for the solar-powered car to optimize its energy production and efficiency. Thus, this research focuses on the analysis of solar irradiance at the EPIC Solar Farm pathway in Kemaman, Terengganu, and investigates the impact on the performance of a solar-powered car.

This article starts with an overview of the methodology used to model solar irradiance and outlines its impact on the performance of the solar-powered car. Next, the results highlighting the relationship between solar irradiance levels and the car's performance are presented with a discussion for developing more effective and sustainable solar-powered vehicles. Last, the conclusion emphasizes the significance of accurate solar irradiance predictions and its potential to enhance solar-powered cars' overall efficiency and performance.

2 Materials and Method

Solar irradiance is the amount of solar energy that falls on a given surface per unit area per unit time [17, 18]. Accurate measurement of solar irradiance is crucial in the design and optimization of solar-powered cars. Spectral irradiance is the distribution of solar irradiance over a range of wavelengths. It is measured in units of watts per square meter per nanometer ($W/m^2/nm$). Spectral irradiance is important in determining the efficiency of solar cells, as different wavelengths of light are absorbed by different types of solar cells. Spectral irradiance is typically measured using a spectroradiometer, which is an instrument that measures the spectral distribution of solar radiation [19].

The irradiance for this research was collected using a pyranometer mounted on the car roof is expressed as Eq. 1.

$$I_{tot} = I_{beam} \cos\theta + I_{diff} \tag{1}$$

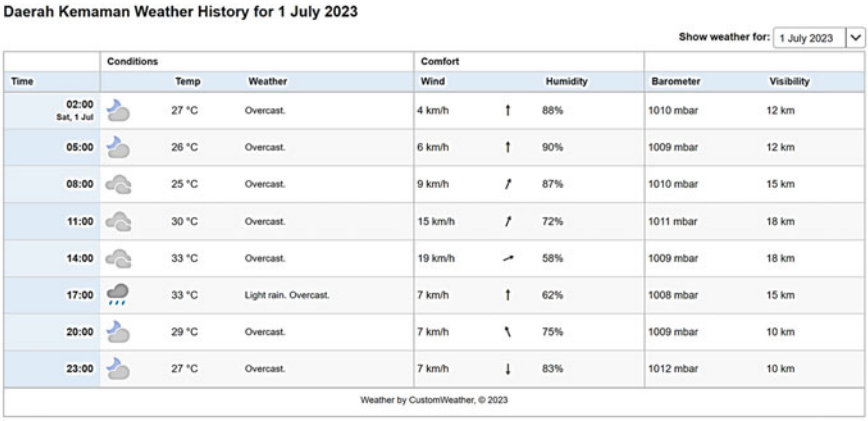
where I_{tot} indicate the total irradiance (W/m^2), I_{beam} intensity of the light from the beam solar, θ indicate angle between the incident ray and the normal to the horizontal instrument plane and I_{diff} indicate intensity of the light from diffuse solar.

The data collection for this research combined primary and secondary sources. The primary data was directly from the experiment, while the secondary data, such as weather, sunrise, and sunset time, was used to validate the solar irradiation analysis retrieved from the real-world data provided by the Time and Date AS. Table 1 summarizes the details specification of the data retrieved. The weather overview is shown in Fig. 1, while Fig. 2 shows the sunrise and sunset time data.

This research was conducted at EPIC Solar Farm Sdn. Bhd, one of Malaysia's largest solar farms, has a total capacity of 50 MWp [22], as shown in Fig. 3. The site is situated in the Kemaman district of Terengganu, near highways, and has an existing charging infrastructure, making it an ideal location for deploying solar-powered cars. The selection criteria included factors such as solar irradiance levels, population density, and availability of charging infrastructure.

Table 1 Specification datasets

Data set	Source	Date	Location
Weather	timeanddate.com [20]	1 July 2023	Kemaman
Sunrise and sunset time	timeanddate.com [21]	1 July 2023	Kemaman



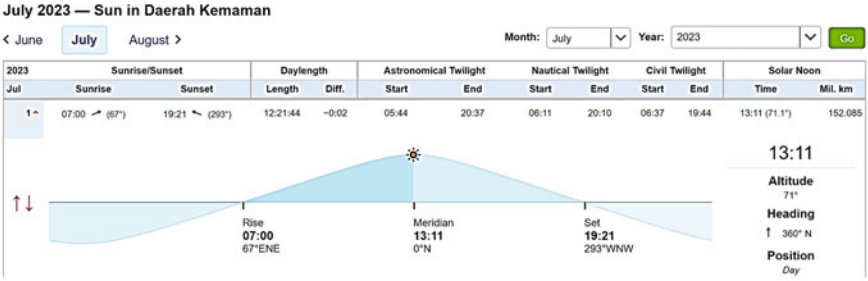
Daerah Kemaman Weather History for 1 July 2023

Show weather for: 1 July 2023

Time	Conditions	Temp	Weather	Comfort Wind		Humidity	Barometer	Visibility
02:00 Sat, 1 Jul		27 °C	Overcast.	4 km/h	↑	88%	1010 mbar	12 km
05:00		26 °C	Overcast.	6 km/h	↑	90%	1009 mbar	12 km
08:00		25 °C	Overcast.	9 km/h	↗	87%	1010 mbar	15 km
11:00		30 °C	Overcast.	15 km/h	↗	72%	1011 mbar	18 km
14:00		33 °C	Overcast.	19 km/h	→	58%	1009 mbar	18 km
17:00		33 °C	Light rain. Overcast.	7 km/h	↑	62%	1008 mbar	15 km
20:00		29 °C	Overcast.	7 km/h	↖	75%	1009 mbar	10 km
23:00		27 °C	Overcast.	7 km/h	↓	83%	1012 mbar	10 km

Weather by CustomWeather, © 2023

Fig. 1 Weather data



July 2023 — Sun in Daerah Kemaman

< June | July | August >

Month: July Year: 2023 Go

2023 Jul	Sunrise/Sunset Sunrise	Sunset	Daylength Length	Diff.	Astronomical Twilight Start	End	Nautical Twilight Start	End	Civil Twilight Start	End	Solar Noon Time	Mil. km
1	07:00 (67°)	19:21 (293°)	12:21:44	−0:02	05:44	20:37	06:11	20:10	06:37	19:44	13:11 (71.1°)	152.085

Rise 07:00 67°ENE

Meridian 13:11 0°N

Set 19:21 293°WNW

13:11

Altitude 71°

Heading ↑ 360° N

Position Day

Fig. 2 Sunset and sunrise time

3 Result

The solar irradiance for 24 h on the first day was recorded, as shown in Fig. 4, with the amount of irradiance recorded over time. The consistent patterns of solar irradiance were recorded from 0.01 a.m. to 8.00 a.m. However, it started to increase from 8.00 a.m. to 12.30 p.m. The decrement pattern was shown from 12.30 p.m. and again constant from 6.30 p.m. to 11.59 p.m. Therefore, the illuminance was constant from 4.30 p.m. to 5.30 p.m. but increased at 6.00 a.m.

Figure 5 shows the solar irradiance recorded from the second day. The solar irradiance was constant from 0.01 a.m. to 8.00 a.m., then increased from 8.00 a.m. to 12.30 p.m. The decrement started at 12.30 p.m. and began constantly from 7.00 to 11.59 p.m.

Tables 2 and 3 retrieved the real-world weather data in Kemaman, Terengganu, while Table 4 shows the sunrise and sunset time. This data was compared to the solar irradiation in Figs. 4 and 5.



Fig. 3 EPIC solar farm Sdn. Bhd

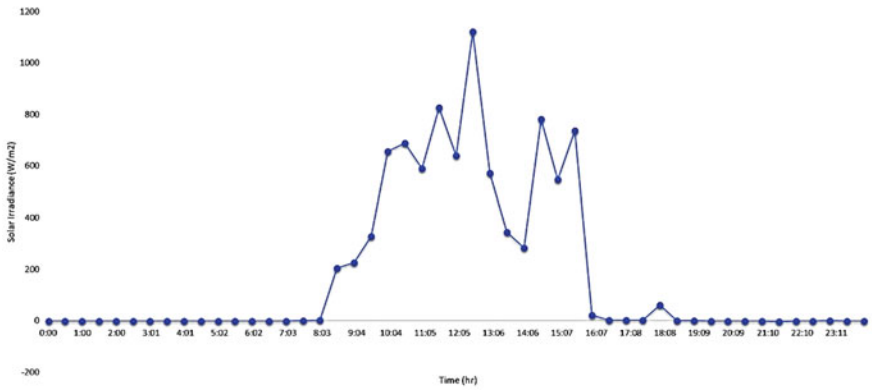


Fig. 4 Solar irradiance on the first day

The solar output voltage recorded on the first day increased from 7.00 a.m. and decreased starting at 4.30 p.m. Figure 6 shows that the solar voltage output is minimum during the period.

Similarly, the solar output voltage for the second day increased from 6.00 a.m. and decreased at 7.00 p.m. Apart from the period, the output voltage is minimum (Fig. 7).

The results show that the relationship between solar irradiance and output voltage is proportional, as verified by the weather and sun time dataset. The mean solar irradiance during daytime does not exceed 1000 for the first and second days because the

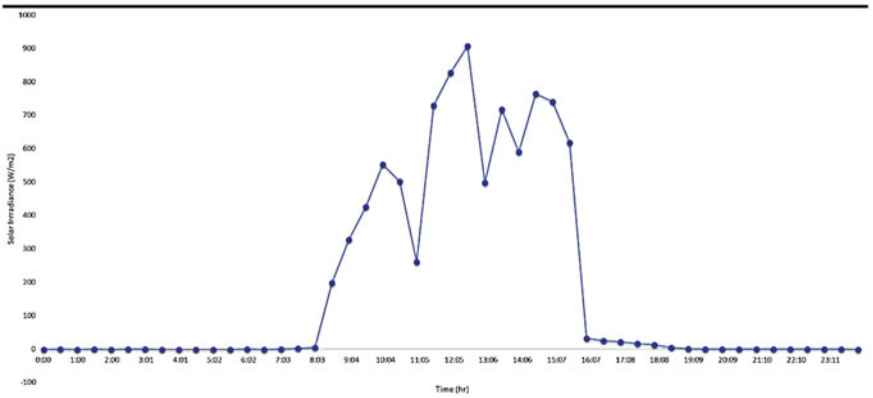


Fig. 5 Solar Irradiance for the second day

Table 2 The weather for the first day

Time	Weather	Temperature (°C)
2.00	Overcast	27
5.00	Overcast	26
8.00	Overcast	25
11.00	Overcast	30
14.00	Overcast	33
17.00	Light rain, overcast	33
20.00	Overcast	29
23.00	Overcast	27

Table 3 The weather for the second day

Time	Weather	Temperature (°C)
2.00	Thundershowers, overcast	26
5.00	Thundershowers, overcast	25
8.00	Low clouds	25
11.00	Overcast	31
14.00	Overcast	33
17.00	Overcast	32
20.00	Overcast	30
23.00	Overcast	28

Table 4 The sunset and sunrise time

Day	Sunrise	Solar noon	Sunset
1	7.00	13.11	19.21
2	7.00	13.11	19.22

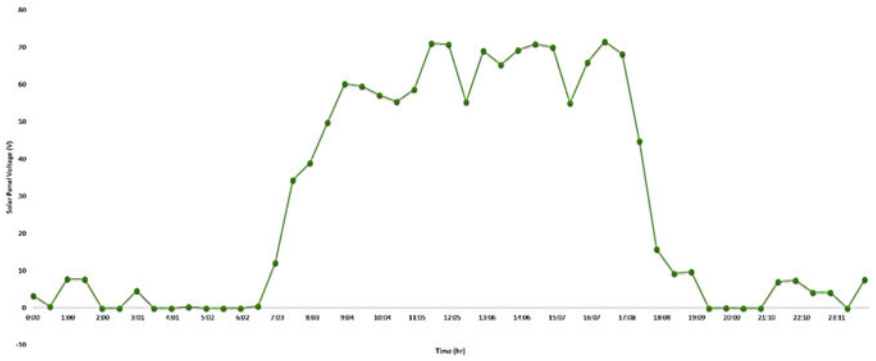


Fig. 6 Solar output voltage for the first day

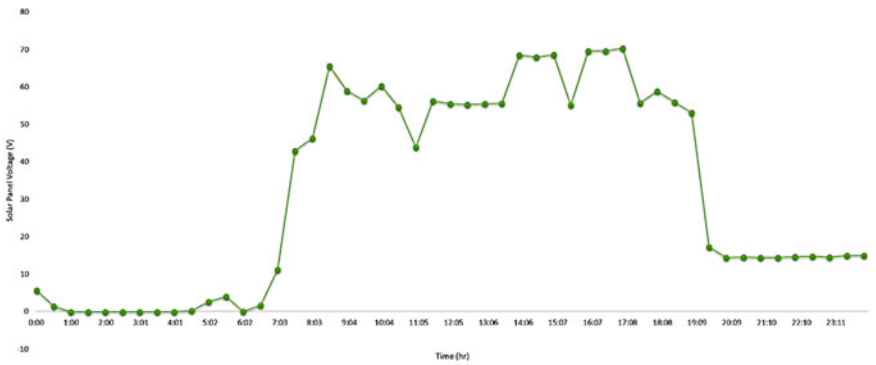


Fig. 7 Solar output voltage for the second day

weather was overcast. The World Meteorological Organization defined that overcast is the meteorological condition of clouds obscuring at least 95% of the sky. The mean solar irradiance for the first day was 408.15; the highest was 1121.46. For the second day, the mean solar irradiance was 416.28, with the highest of 906.74. Thus, it proved that the high reading of solar irradiance is around noon. Besides that, rainy days also affected the solar irradiance reading. On the first day at 5.00 p.m., the weather was light rain; thus, the irradiance below 100 was recorded. Thunderstorm weather conditions on the second day, in the early morning, around 2.00 a.m. to 5.00 a.m., recorded the low reading irradiance during the sunset.

4 Discussion

The results are the evidence that solar irradiance plays a significant role in the performance of the solar-powered car. The positive correlation between solar irradiance and the voltage output of the solar panels indicates that higher solar irradiance leads to increased power generation and improved car performance. This finding aligned with the basic understanding that solar panels use sunlight to generate electricity. The solar panel efficiency for the first day was 7.6%, while 6.4% for the second day. The results validated the importance of accurate solar irradiance predictions for optimizing the performance and efficiency of solar-powered cars.

The existing research focused on the impact of solar irradiance on solar-powered vehicles generally. However, this research contributes to the existing body of knowledge by focusing on the specific context of the EPIC Solar Farm pathway in Kemaman, Terengganu. Considering the solar irradiance levels at this location, the findings provide location-specific insights that can feed the design and development of solar-powered cars for similar environments. Additionally, this research underscores the significance of considering solar irradiance variations in optimizing the efficiency and range of solar-powered vehicles.

Despite the valuable findings and insights gained from this research, it is essential to acknowledge its limitations. One potential weakness is the specific focus on the EPIC Solar Farm pathway in Kemaman, Terengganu. While this allows for a detailed analysis of solar irradiance in a particular location, the generalizability of the findings to other regions or environments may be limited. Future research could benefit from broader geographical coverage to address the limitation.

The implications of this research extend beyond the specific context of the EPIC Solar Farm pathway. By emphasizing the importance of accurate solar irradiance predictions and its direct influence on the performance and efficiency of solar-powered cars, this study has implications for the broader field of solar-powered transportation. Car manufacturers and designers can leverage solar irradiance modeling to optimize the design and performance of solar-powered vehicles in various geographical locations. Implementing location-specific solar irradiance predictions can enhance solar-powered cars' overall efficiency, range, and reliability, thereby promoting the adoption of sustainable and clean energy solutions in the transportation sector.

5 Conclusion

This study aimed to investigate the impact of solar irradiance on the efficiency and performance of solar-powered cars, explicitly focusing on the EPIC Solar Farm pathway in Kemaman, Terengganu. By analyzing the relationship between solar irradiance levels and the voltage output of solar panels installed on the car, the study aimed to enhance the overall efficiency and range of solar-powered vehicles.

The findings of this study indicate a clear positive correlation between solar irradiance levels and the performance of the solar-powered car. As solar irradiance increased, the voltage output of the solar panels also rose, resulting in improved power generation and performance. The higher solar panel energy efficiency recorded is 7.6%. Solar irradiance and weather are not the only factor affecting the solar panel's performance. Otherwise, load resistance, cell temperature [23, 24], shading [25, 26], soiling [27], solar charge controller [28, 29], and maintenance service [30, 31] are also involved.

Future works should consider expanding the geographical coverage to include multiple locations and diverse environmental conditions to build upon this research. Additionally, incorporating a larger sample size and employing more comprehensive data collection techniques would enhance the accuracy and reliability of the findings. Further investigations can explore advanced modeling techniques to predict solar irradiance levels at specific locations, facilitating the optimization of solar-powered vehicle designs and maximizing their efficiency and range.

References

1. Sharif A, Meo MS, Chowdhury MAF, Sohag K (2021) Role of solar energy in reducing ecological footprints: an empirical analysis. J Clean Prod 292:126028. https://doi.org/10.1016/J.JCLEPRO.2021.126028
2. Waseem M, Sherwani AF, Suhaib M (2019) Integration of solar energy in electrical, hybrid, autonomous vehicles: a technological review. https://doi.org/10.1007/s42452-019-1458-4
3. Oh M, Kim SM, Park HD (2020) Estimation of photovoltaic potential of solar bus in an urban area: case study in Gwanak, Seoul, Korea. Renew Energy 160:1335–1348. https://doi.org/10.1016/j.renene.2020.07.048
4. Diahovchenko I, Petrichenko L, Borzenkov I, Kolcun M (2022) Application of photovoltaic panels in electric vehicles to enhance the range. Heliyon 8. https://doi.org/10.1016/j.heliyon.2022.e12425
5. Kaldellis JK, ZD, BG, K (2019) Performance analysis of a small solar-powered electric vehicle in Greece. Energies (Basel)
6. Zhu R, Kondor D, Cheng C, Zhang X, Santi P, Wong MS, Ratti C (2022) Solar photovoltaic generation for charging shared electric scooters. Appl Energy 313:118728. https://doi.org/10.1016/J.APENERGY.2022.118728
7. Thiel C, Gracia Amillo A, Tansini A, Tsakalidis A, Fontaras G, Dunlop E, Taylor N, Jäger-Waldau A, Araki K, Nishioka K, Ota Y, Yamaguchi M (2022) Impact of climatic conditions on prospects for integrated photovoltaics in electric vehicles. Renew Sustain Energy Rev 158. https://doi.org/10.1016/j.rser.2022.112109
8. Mohammad A, Zamora R, Lie TT (2020) Integration of electric vehicles in the distribution network: a review of PV based electric vehicle modelling. Energies (Basel) 13. https://doi.org/10.3390/en13174541
9. Fachrizal R, Shepero M, van der Meer D, Munkhammar J, Widén J (2020) Smart charging of electric vehicles considering photovoltaic power production and electricity consumption: a review. eTransportation 4:100056. https://doi.org/10.1016/J.ETRAN.2020.100056
10. Higashitani T, Ikegami T, Uemichi A, Akisawa A (2021) Evaluation of residential power supply by photovoltaics and electric vehicles. Renew Energy 178:745–756. https://doi.org/10.1016/J.RENENE.2021.06.097

11. Sunddararaj SP, Rangarajan SS, Nallusamy S, Randolph Collins E, Senjyu T (2021). A brief survey on important interconnection standards for photovoltaic systems and electric vehicles. https://doi.org/10.3390/wevj12030117
12. Sagaria S, Duarte G, Neves D, Baptista P (2022) Photovoltaic integrated electric vehicles: assessment of synergies between solar energy, vehicle types and usage patterns. J Clean Prod 348:131402. https://doi.org/10.1016/J.JCLEPRO.2022.131402
13. Tawa H, Saiki H, Ota Y, Araki K, Takamoto T, Nishioka K (2020) Accurate output forecasting method for various photovoltaic modules considering incident angle and spectral change owing to atmospheric parameters and cloud conditions. https://doi.org/10.3390/app10020703
14. Araki K, Ota Y, Yamaguchi M (2020) Measurement and modeling of 3D solar irradiance for vehicle-integrated photovoltaic. Appl Sci (Switzerland) 10. https://doi.org/10.3390/app10030872
15. Park C, Park H, Jeon H, Choi K, Suh J (2023) Evaluation and validation of photovoltaic potential based on time and pathway of solar-powered electric vehicle. Appl Sci 13:1025. https://doi.org/10.3390/APP13021025
16. Sierra Rodriguez A, de Santana T, MacGill I, Ekins-Daukes NJ, Reinders A (2020) A feasibility study of solar PV-powered electric cars using an interdisciplinary modeling approach for the electricity balance, CO_2 emissions, and economic aspects: the cases of The Netherlands, Norway, Brazil, and Australia. Prog Photovoltaics Res Appl 28:517–532. https://doi.org/10.1002/pip.3202
17. Kherbiche Y, Ihaddadene N, Ihaddadene R, Hadji F, Mohamed J, Beghidja AH (2021) Solar energy potential evaluation. Case of study: M'Sila, an Algerian Province. Int J Sustain Dev Plann 16:1501–1508. https://doi.org/10.18280/ijsdp.160811
18. Kumari P, Toshniwal D (2021) Deep learning models for solar irradiance forecasting: a comprehensive review. J Clean Prod 318:128566. https://doi.org/10.1016/J.JCLEPRO.2021.128566
19. Polo J, Alonso-Abella M, Martín-Chivelet N, Alonso-Montesinos J, López G, Marzo A, Nofuentes G, Vela-Barrionuevo N (2020) Typical meteorological year methodologies applied to solar spectral irradiance for PV applications. Energy 190:116453. https://doi.org/10.1016/J.ENERGY.2019.116453
20. Weather in July 2023 in Daerah Kemaman, Malaysia. https://www.timeanddate.com/weather/@1732963/historic?month=7&year=2023. Last accessed 29 July 2023
21. Sunrise and sunset times in Daerah Kemaman, July 2023. https://www.timeanddate.com/sun/@1732963?month=7&year=2023. Last accessed 29 July 2023
22. EPIC Group - Subsidiaries | EPIC Solar Sdn Bhd. https://www.epicgroup.com.my/subsidiaries/epic-solar.php. Last accessed 29 July 2023
23. Mustafa RJ, Gomaa MR, Al-Dhaifallah M, Rezk H (2020) Environmental impacts on the performance of solar photovoltaic systems. Sustainability (Switzerland) 12. https://doi.org/10.3390/su12020608
24. Dwivedi P, Sudhakar K, Soni A, Solomin E, Kirpichnikova I (2020) Advanced cooling techniques of P.V. modules: a state of art. Case Stud Thermal Eng 21: 100674. https://doi.org/10.1016/J.CSITE.2020.100674
25. Wang D, Qi T, Liu Y, Wang Y, Fan J, Wang Y, Du H (2020) A method for evaluating both shading and power generation effects of rooftop solar PV panels for different climate zones of China. Sol Energy 205:432–445. https://doi.org/10.1016/J.SOLENER.2020.05.009
26. Kunz O, Evans RJ, Juhl MK, Trupke T (2020) Understanding partial shading effects in shingled PV modules. Sol Energy 202:420–428. https://doi.org/10.1016/J.SOLENER.2020.03.032
27. Salamah T, Ramahi A, Alamara K, Juaidi A, Abdallah R, Abdelkareem MA, Amer EC, Olabi AG (2022) Effect of dust and methods of cleaning on the performance of solar PV module for different climate regions: comprehensive review. Sci Total Environ 827:154050. https://doi.org/10.1016/J.SCITOTENV.2022.154050
28. Yap KY, Sarimuthu CR, Lim JMY (2020) Artificial intelligence based MPPT techniques for solar power system: a review. J Mod Power Syst Clean Energy 8:1043–1059. https://doi.org/10.35833/MPCE.2020.000159

29. Katche ML, Makokha AB, Zachary SO, Adaramola MS (2023) A comprehensive review of maximum power point tracking (MPPT) techniques used in solar PV systems. https://doi.org/10.3390/en16052206
30. Aboagye B, Gyamfi S, Ofosu EA, Djordjevic S (2022) Investigation into the impacts of design, installation, operation and maintenance issues on performance and degradation of installed solar photovoltaic (PV) systems. Energy Sustain Dev 66:165–176. https://doi.org/10.1016/J.ESD.2021.12.003
31. Hasan K, Yousuf SB, Tushar MSHK, Das BK, Das P, Islam MS (2022). Effects of different environmental and operational factors on the PV performance: a comprehensive review. https://doi.org/10.1002/ese3.1043

Revolutionizing Transportation: Analyzing Solar Car Efficiency at EPIC Solar Farm



Afidatul Nadia Mok Hat, Ruzlaini Ghoni, Mohd Tarmizi Ibrahim, Ahmad Firdaus Zali, and Fuaad Mohamed Nawawi

Abstract The transportation sector's heavy reliance on fossil fuels significantly contributes to air pollution and climate change. There is a growing interest in using solar-powered cars to address these environmental challenges and promote sustainable transportation. However, there is a need to comprehensively explore the performance of solar-powered cars to understand their efficiency and effectiveness. The primary objective of this research is to comprehensively explore the performance of solar-powered cars along the EPIC Solar Farm pathway in Kemaman, Terengganu. The development comprised an array of sensors, communication modules, and onboard processing systems that were integrated into the cars to facilitate the real-time collection, interpretation, and transmission of crucial data. This dataset encompasses vital information relating to battery consumption, environmental factors, and navigation elements. The findings of this study underscore a distinct relationship between solar energy and battery energy. Specifically, as the conversion of solar radiation into solar energy intensifies, the stored energy within the battery increases proportionally. Moreover, the study places significant emphasis on the influence of battery usage on overall performance, particularly during operation along the EPIC Solar Farm pathway. Notably, the study investigates the impact of charging and discharging capacity and battery usage to understand the buggy's efficiency and effectiveness comprehensively. In summary, the results of this study highlight the substantial benefits associated with the adoption of solar energy in transportation, particularly within the context of solar farm environments. The findings underscore the positive implications for air quality improvement, climate change mitigation, and promoting sustainable practices within the transportation sector.

Keywords Solar-powered car · Renewable energy · Sustainable transportation

A. N. M. Hat (✉) · R. Ghoni · M. T. Ibrahim
Faculty of Engineering Technology, University College TATI, Chukai, Malaysia
e-mail: afidatulnadia@uctati.edu.my

A. F. Zali · F. M. Nawawi
EPIC Solar Farm Sdn. Bhd., Kemaman Supply Base, Chukai, Malaysia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_7

1 Introduction

Transportation is a critical sector that significantly impacts air pollution and contributes to climate change through its heavy reliance on fossil fuels [1–3]. The quest for sustainable transportation solutions becomes more crucial as the world faces increasing environmental challenges. According to the International Energy Agency, transportation is one of the top three sectors with rising CO_2 emissions, behind power at 42% and industry at 26%. The CO_2 emission from transport grew by 254 Mt from 2021 to 2022 [4]. In Malaysia, the combustion of fossil fuels by transportation is one of the significant sources of air pollution [5–8]. Replacing fossil fuels in the transport sector with renewable energy sources is one way to mitigate climate change. Among the promising alternatives, solar-powered cars have emerged as a viable option due to their ability to operate primarily or exclusively on solar energy, thereby eliminating harmful emissions like carbon dioxide. Embracing solar energy in transportation holds excellent potential for mitigating the adverse effects of traditional fuel-based vehicles [9, 10]. While solar-powered cars offer a promising solution, their performance and effectiveness, particularly in the unique context of solar farm environments, remain areas of exploration that require comprehensive investigation.

Scholars have recently delved into solar-powered cars to understand their feasibility and potential benefits. Studies have focused on various aspects, such as solar energy conversion efficiency [11, 12], battery technologies [13, 14], and applications in transportation. Centeno et al. [15] and Oh et al. [16] conducted a study on the design and optimization of solar-powered vehicles for urban commuting, while Hossain et al. [17] explored the influence of load on batter capacities on the range and endurance of solar cars in long-distance travel. Additionally, the researchers also investigated the integration of solar panels with electric vehicles for partial solar supplementation [18–23]. However, while these studies contribute valuable insights, a comprehensive exploration of solar-powered cars' performance along specific demographics and their comparison with conventional buggies remains relatively unexplored.

Despite the existing research efforts in solar-powered cars, there is a noticeable research gap concerning their performance within the specific context of geographical area. Understanding the relationship between solar energy and battery energy, assessing the battery's capacity for sustained trips, and comparing the performance metrics with conventional buggies are crucial aspects that warrant closer investigation. By addressing these gaps, researchers can gain a more comprehensive understanding of the efficiency and effectiveness of solar-powered cars, particularly in solar farm settings.

This article aims to fill the research gap by presenting a case study of solar-powered buggy performance at the EPIC Solar Farm, Kemaman Terengganu. The primary objective of this research is to comprehensively explore the performance of solar-powered buggies operating along the EPIC Solar Farm pathway. An array of sensors, communication modules, and onboard processing systems were integrated into the buggy, enabling real-time data collection, interpretation, and transmission.

The collected dataset encompasses vital information related to battery consumption, environmental factors, and navigation elements. The study focuses on understanding the relationship between solar energy and battery energy, evaluating the battery's capacity for sustained trips, and comparing the performance metrics during day and night. The findings of this research not only contribute to the existing body of knowledge on solar-powered cars but offer insights into the advantages of adopting solar energy in transportation, especially in solar farm environments. The results underscore the positive implications for air quality improvement, climate change mitigation, and promoting sustainable practices within the transportation sector.

2 Materials and Method

Data for this research was collected from solar-powered buggy operating along the EPIC Solar Farm pathway in Kemaman, Terengganu [24]. The real-time data collection was realized by integrating sensors, communication modules, and onboard processing systems into the buggy, as shown in Fig. 1. The sensors captured essential information related to battery consumption, environmental factors, and navigation elements. The data collection tools included data loggers to record and store the sensor readings. Additionally, GPS devices were used for navigation data, and environmental sensors measured factors like temperature and solar radiation. Table 1 summarizes details of items and specifications of a solar-powered buggy.

The type of car that was used in this research is the buggy. This buggy is used by the engineers or technicians at the solar farm for maintenance tasks like cleaning, lubrication, adjustments, repairs, and replacement of the solar farm. Lightweight and ability to conform to the shape of the buggy roof make the buggy car suitable for solar energy implementation. Table 2 shows the general specification of the buggy.

The process of data collection is summarized in Fig. 2. The buggy's performance runs smoothly with solely solar energy sources and powered sensors for real-time data monitoring system and data logging via a cloud platform integrated into the buggy. The data analysis was performed via Microsoft Excel. The collected data were subjected to statistical analysis to draw meaningful conclusions. Descriptive statistics such as mean and standard deviation (SD) were calculated to summarize the data and measure the central tendency and variability. The mean represents the average value of a variable, while the standard deviation indicates the dispersion of data points around the mean.

3 Result

The voltage and current variation for solar panels depend on the solar irradiance, as shown in Fig. 3. The voltage and current increased rapidly during sunrises and decreased during sunset. Solar panels' voltage slightly changes during the movement

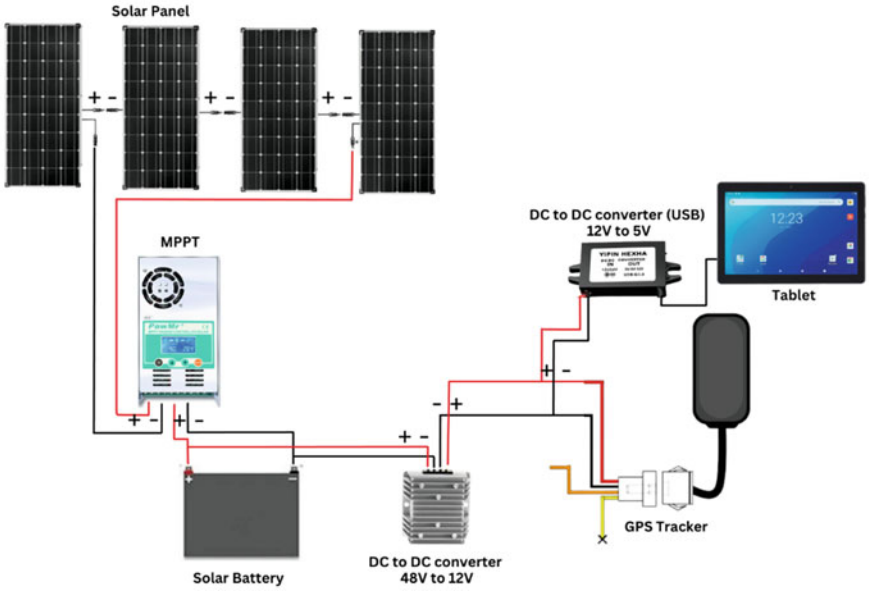


Fig. 1 Schematic diagram of solar-powered buggy

Table 1 Details items and specifications of solar-powered buggy

Item	Description
Solar panel	150 W × 4, Monocrystalline
Solar charge controller	48 V, maximum power point tracking (MPPT)
Solar battery	48 V 50 AH, lithium-ion battery
DC to DC converter	48–12 V and 12–5 V
Tablet	5 V/2 A, 12-inch, Android
GPS tracker	12 V, 8 × 3 × 1 cm, Android

Table 2 Buggy general specification

Item	Description
Dimension	2400 * 1200 * 1800 mm
Engine	48 V
Horsepower	3.3 hp (2.4 kW) DC motor
Battery	51.2 V 80 Ah
Exterior color	Metallic jade green
Weight	470 kg
Loading	2 people
Brakes	Self-adjusting, rear wheel mechanical drum

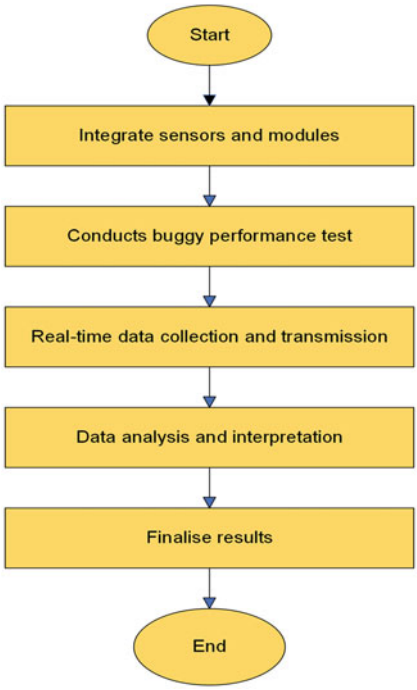


Fig. 2 Experimental procedures

of clouds covering the sun. The intensity of the solar irradiance influences the current, while the increases in the temperature reduce the voltage.

The voltage of the battery increased during the daytime between sunset and sunrise. It is because of the battery in the charging process and the insignificant charging amount recorded when the buggy was used, as shown in Fig. 4. The voltage

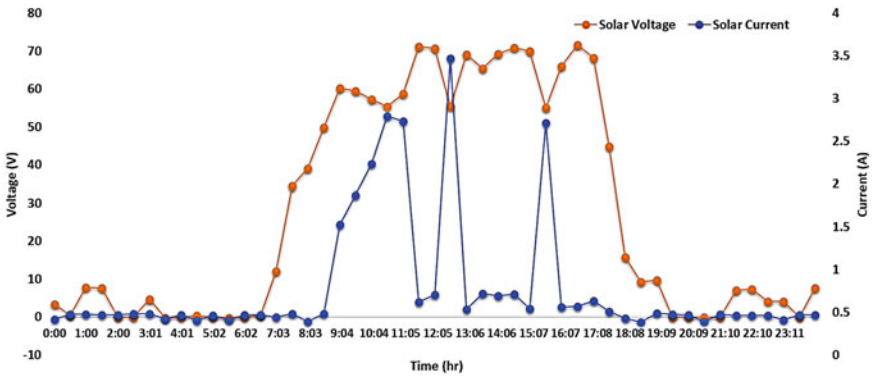


Fig. 3 Voltage and current for solar panel

for the battery decreases and increases as the battery is charging and discharging during the daytime. The battery's current is constant when the buggy car is static and increases when the buggy car is moving. The buggy car used a DC motor, which converts electrical energy into mechanical energy [25]. The DC motor connected to the solar battery converts the direct current into mechanical energy while the buggy moves.

Table 3 shows the mean and standard deviation for the performance of the buggy car during the day and night. The mean and standard deviation of solar irradiance during the day are higher than at night due to irradiance from the entire disk of the Sun and almost constant during the night. The mean and the standard deviation of solar voltage, current, and battery voltage depend on the solar irradiance during the daytime while it is in charging mode. At night, the standard deviation of the battery voltage is lower, with the battery voltage almost constant because the battery is on standby mode, and there is no drawback current due to the solar charger controller. Maximum power point tracking (MPPT) was used to regulate the charge of the solar battery.

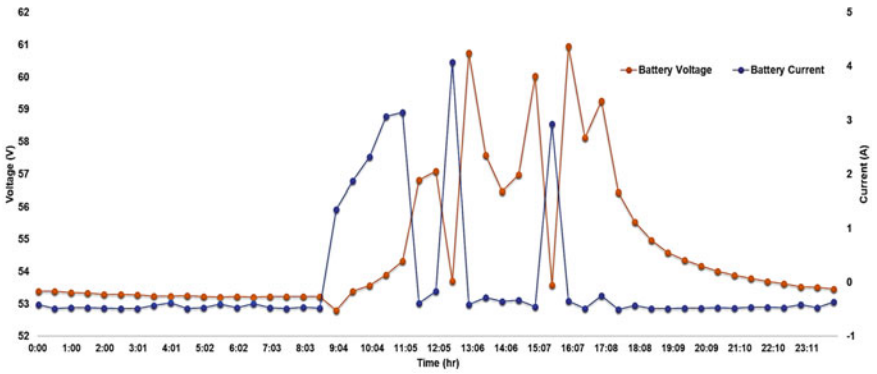


Fig. 4 Voltage and current for solar battery

Table 3 Mean value

Item	Mean (μ)		Standard deviation (σ)	
	Day	Night	Day	Night
Solar irradiance	408.15	1.44	329.70	11.63
Solar voltage	58.60	4.80	14.52	8.77
Solar current	1.21	0.46	0.96	0.03
Battery voltage	55.82	53.77	2.72	0.77
Battery current	0.64	−0.46	1.52	0.04
Battery percentage	96.14	94.96	6.61	3.96

Table 4 summarizes the charging rate of the solar-powered buggy during morning, noon, and evening. The charging rate is fast during the noon with higher mean solar irradiance followed by evening and morning. Solar irradiance is vital to a solar-powered car's charging rate.

Figure 5 shows the battery capacity and day ambient temperature. The fastest decrease in the battery capacity was recorded at 9.00 a.m. due to its movement. The battery capacity percentage is maintained at 90% when the temperature is between 20 and 40 °C. The battery capacity shows insignificant changes when the buggy is not moving, proving the efficient battery performance. The lithium-ion battery was used as storage, highlighted in the datasheet as efficient at low temperatures (Fig. 6).

This solar-powered buggy is solely powered by solar energy and not fed by fossil fuel or electric charging. This solar-powered buggy produces zero carbon emissions compared to petrol cars, producing 2.296 kg of carbon emission per liter, while diesel will produce 2.653 kg per liter [27].

Table 4 The battery charging rate

Battery voltage (V)	Battery percentage (%)	Time (min)	Mean solar irradiance
1.19	10	102	534.7198571
1.53	26	30	762.087129
1.52	21	32	717.7227813

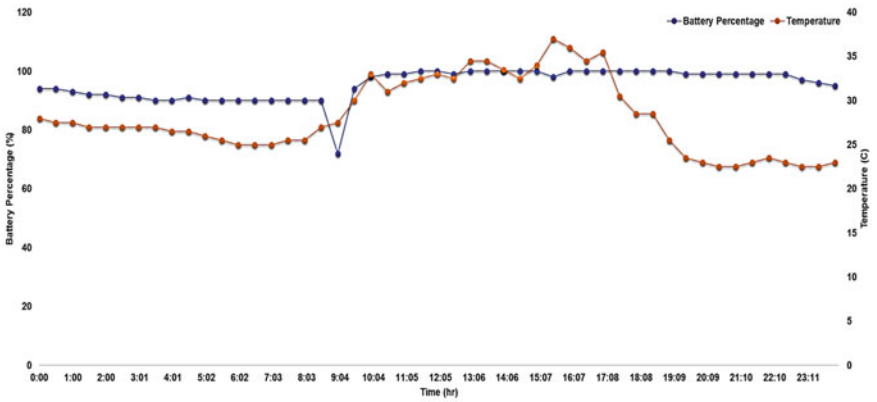


Fig. 5 Battery percentage and ambient temperature

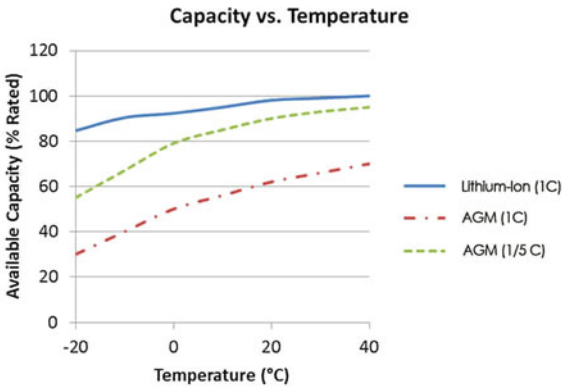


Fig. 6 Datasheet of solar battery [26]

4 Discussion

This study offers valuable insights into the performance of solar-powered buggy operating along the EPIC Solar Farm pathway. The analysis revealed a distinct relationship between solar and battery energy, with an increase in solar energy conversion leading to a proportional rise in the stored energy within the battery. This finding highlights the effectiveness of solar-powered cars in harnessing solar radiation and converting it into usable energy for propulsion. The study also emphasized the significant influence of battery usage on overall performance, particularly during operation along the solar farm pathway. The ability of the battery to sustain multiple trips within a day demonstrates the endurance and reliability of solar-powered cars. Overall, these results underscore solar-powered cars' feasibility and potential benefits in mitigating air pollution, combating climate change, and promoting sustainable practices within the transportation sector.

Compared to related studies, our research contributes novel insights by focusing on solar-powered car performance within a solar farm environment. This unique context provides relevant data for solar farm operators and transportation planners seeking eco-friendly transportation solutions within solar energy generation facilities. Moreover, our research extends beyond solar energy conversion efficiency and battery technologies to encompass comprehensive performance metrics. This comparative analysis with conventional buggies further strengthens the value of our findings, offering practical insights into the advantages of adopting solar energy in transportation.

Despite the valuable contributions and insights, our study has certain limitations. First, the research was conducted at a specific solar farm site in Kemaman, Terengganu, which may limit the generalizability of the results to other solar farm environments with different geographical and environmental conditions. Additionally, the data collection period may not encompass all possible variations in weather conditions, potentially influencing the performance metrics observed.

This research holds significant implications for various transportation and renewable energy stakeholders. The findings highlight the substantial benefits of adopting solar energy in transportation, especially in solar farm environments. Solar-powered cars offer a promising solution for mitigating air pollution and combating climate change, aligning with global efforts to promote sustainable practices in the transportation sector. The demonstrated relationship between solar energy and battery energy underscores the potential for further advancements in solar energy technologies to enhance the efficiency and performance of solar-powered vehicles. Transportation planners and policymakers can utilize these insights to integrate solar-powered buggies into their fleet, particularly within solar farm facilities, to achieve eco-friendly transportation and promote renewable energy utilization.

5 Conclusion

The primary objective of this research was to comprehensively explore the performance of solar-powered buggy along the EPIC Solar Farm pathway in Kemaman, Terengganu. While this study has provided valuable insights into the performance of solar-powered buggies at the EPIC Solar Farm, several avenues for future research and development exist. Firstly, extending the research to multiple solar farm sites with varying environmental conditions would enhance the generalizability of the findings. Further investigations into the impacts of different solar energy conversion technologies and battery types on overall performance could provide a deeper understanding of solar-powered buggy efficiency. Long-term studies examining the buggy performance over various seasons and weather conditions would strengthen the robustness of the results.

Additionally, exploring potential hybrid systems that combine solar energy with other renewable sources or energy storage technologies could offer more sustainable transportation solutions. Lastly, integrating feedback from users and operators of solar-powered buggies to assess practical challenges and user experiences would help refine the implementation and promote the adoption of solar-powered transportation. By addressing these areas, future research can build upon the foundation laid by this research and contribute to advancing eco-friendly and efficient solar-powered transportation systems.

References

1. Xue Y, Cao X, Ai Y, Xu K, Zhang Y (2020) Primary air pollutants emissions variation characteristics and future control strategies for transportation sector in Beijing, China. Sustainability (Switzerland) 12. https://doi.org/10.3390/su12104111
2. Gallo M, Marinelli M (2020). Sustainable mobility: a review of possible actions and policies. https://doi.org/10.3390/su12187499

3. Adams S, Boateng E, Acheampong AO (2020) Transport energy consumption and environmental quality: does urbanization matter? Sci Total Environ 744:140617. https://doi.org/10.1016/J.SCITOTENV.2020.140617
4. Energy agency, international: CO_2 Emissions in 2022
5. Mohd Shafie SH, Mahmud M, Mohamad S, Rameli NLF, Abdullah R, Mohamed AF (2022) Influence of urban air pollution on the population in the Klang Valley, Malaysia: a spatial approach. Ecol Process 11. https://doi.org/10.1186/s13717-021-00342-0
6. Solaymani S (2022) CO_2 emissions and the transport sector in Malaysia. Front Environ Sci 9. https://doi.org/10.3389/fenvs.2021.774164
7. Mohd Shafie SH, Mahmud M (2020) Urban air pollutant from motor vehicle emissions in Kuala Lumpur, Malaysia. Aerosol Air Qual Res 20:2793–2804. https://doi.org/10.4209/aaqr.2020.02.0074
8. Brohi SN, Pillai TR, Asirvatham D, Ludlow D, Bushell J (2018) Towards smart cities development: a study of public transport system and traffic-related air pollutants in Malaysia. In: IOP conference series: earth and environmental science. Institute of Physics Publishing. https://doi.org/10.1088/1755-1315/167/1/012015
9. Holechek JL, Geli HME, Sawalhah MN, Valdez R (2022) A global assessment: can renewable energy replace Fossil fuels by 2050? Sustainability (Switzerland) 14. https://doi.org/10.3390/su14084792
10. Kou G, Yüksel S, Dinçer H (2022) Inventive problem-solving map of innovative carbon emission strategies for solar energy-based transportation investment projects. Appl Energy 311. https://doi.org/10.1016/j.apenergy.2022.118680
11. Chiu JS, Zhao YM, Zhang S, Wuu DS (2020) The role of laser ablated backside contact pattern in efficiency improvement of mono crystalline silicon PERC solar cells. Sol Energy 196:462–467. https://doi.org/10.1016/J.SOLENER.2019.12.044
12. Zhang Y, Ren J, Pu Y, Wang P (2020) Solar energy potential assessment: a framework to integrate geographic, technological, and economic indices for a potential analysis. Renew Energy 149:577–586. https://doi.org/10.1016/J.RENENE.2019.12.071
13. Chen T, Jin Y, Lv H, Yang A, Liu M, Chen B, Xie Y, Chen Q (2020) Applications of lithium-ion batteries in grid-scale energy storage systems. https://doi.org/10.1007/s12209-020-00236-w
14. Zhang Y, Hao N, Lin X, Nie S (2020) Emerging challenges in the thermal management of cellulose nanofibril-based supercapacitors, lithium-ion batteries and solar cells: a review. Carbohydr Polym 234:115888. https://doi.org/10.1016/J.CARBPOL.2020.115888
15. Centeno Brito M, Santos T, Moura F, Pera D, Rocha J (2021) Urban solar potential for vehicle integrated photovoltaics. Transp Res D Transp Environ 94:102810. https://doi.org/10.1016/J.TRD.2021.102810
16. Oh M, Kim SM, Park HD (2020) Estimation of photovoltaic potential of solar bus in an urban area: case study in Gwanak, Seoul, Korea. Renew Energy 160:1335–1348. https://doi.org/10.1016/j.renene.2020.07.048
17. Hossain MA (2019) Performance investigation of a solar powered electric car
18. Diahovchenko I, Petrichenko L, Borzenkov I, Kolcun M (2022) Application of photovoltaic panels in electric vehicles to enhance the range. Heliyon 8. https://doi.org/10.1016/j.heliyon.2022.e12425
19. Rutten B, Cobbenhagen R (2019) Future trends in electric vehicles enabled by internet connectivity, solar, and battery technology. In: Automotive systems and software engineering. Springer International Publishing, pp 323–346. https://doi.org/10.1007/978-3-030-12157-0_15
20. Waseem M, Sherwani AF, Suhaib M (2019) Integration of solar energy in electrical, hybrid, autonomous vehicles: a technological review. https://doi.org/10.1007/s42452-019-1458-4
21. Sagaria S, Duarte G, Neves D, Baptista P (2022) Photovoltaic integrated electric vehicles: assessment of synergies between solar energy, vehicle types and usage patterns. J Clean Prod 348:131402. https://doi.org/10.1016/J.JCLEPRO.2022.131402
22. Sunddararaj SP, Rangarajan SS, Nallusamy S, Randolph Collins E, Senjyu T (2021). A brief survey on important interconnection standards for photovoltaic systems and electric vehicles. https://doi.org/10.3390/wevj12030117

23. Fachrizal R, Shepero M, van der Meer D, Munkhammar J, Widén J (2020) Smart charging of electric vehicles considering photovoltaic power production and electricity consumption: a review. eTransportation 4:100056. https://doi.org/10.1016/J.ETRAN.2020.100056
24. EPIC Group - Subsidiaries | EPIC Solar Sdn Bhd. https://www.epicgroup.com.my/subsidiaries/epic-solar.php. Last accessed 29 July 2023
25. By E, Reddy KS (2023) Electrical devices and systems
26. Lithium-ion Battery Charging and Advantages—PowerTech Systems. https://www.powertechsystems.eu/home/tech-corner/lithium-ion-battery-advantages/. Last Accessed 03 August 2023
27. Carbon Footprint Calculation—A small introduction of ISO 14064—Green Clean Guide. https://greencleanguide.com/calculate-your-carbon-footprint/. Last Accessed 03 August 2023

Mathematical Modelling of Electromagnetic Transduction for Optimal Power Harvesting from Induction Motors



Ammar Husaini Hussian, Ruzlaini Ghoni, Mohd Tarmizi Ibrahim, Shaiful Rizalmee Wahid, Afidatul Nadia Mok Hat, Mohd Aizat Sulaiman, and Mohd Fadhil Ibrahim

Abstract Induction motors are widely used to convert electrical power into mechanical power because they are simple, rugged, robust, efficient, and suitable for applications in harsh environments. Previous studies have been conducted on magnetic energy harvesting from manufacturing machinery, such as an induction motor, the electromagnetic transducer developed employed the clamped current-transformer technique. Until now, no study has been conducted on mathematical modelling for harvesting magnetic energy using a Clampless Current-Transformer method, specifically from an induction motor. This article proposes the mathematical modelling of electromagnetic transduction from an induction motor using Ansys Engineering Simulation Software. It models and simulates the core design with each material before using Clampless Current-Transformer method electromagnetic transducer to capture the radiation of magnetic field energy in various core designs. The simulation contrasts Stainless and Carbon Steel for the core material consisting of different core shapes; cylindrical and square rods with 300 and 500 turns. The proposed electromagnetic transduction provides better flexibility due to its clampless feature from the incoming supply to the induction motor; instead, it can be placed anywhere near the most substantial magnetic field radiation. The modelling demonstrates that long cylindrical core Carbon Steel produces higher output power for 500 turns. The proposed modelling with the Clampless Current-Transformer approach can efficiently harvest magnetic field energy. It could have practical implications for energy harvesting from induction motors, enabling the transducer to be conveniently near the strongest magnetic field without clamping to the motor supply.

Keywords Electromagnetic harvesting · Clampless current-transformer · Stainless steel and carbon steel · Electromagnetic transducer

A. H. Hussian (✉) · R. Ghoni · M. T. Ibrahim · S. R. Wahid · A. N. M. Hat · M. A. Sulaiman · M. F. Ibrahim
Faculty of Engineering Technology, University College TATI, Chukai, Malaysia
e-mail: ammarhusaini@uctati.edu.my

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_8

1 Introduction

Induction motors, renowned for their versatility, ruggedness, and efficiency, play a fundamental role in converting electrical energy into mechanical power across various applications. This transformative process not only drives machinery but also engenders a low-magnitude magnetic field as a by product a field which has long piqued the interest of researchers due to its latent potential for energy harvesting [1]. Harnessing this magnetic energy as an auxiliary power source could revolutionize energy efficiency and sustainability, addressing the exigent concern of unstable power supplies for sensors and other low-power devices. While previous research has primarily focused on capturing magnetic energy from diverse sources, including transmission lines and power plants, a critical research gap emerges in the domain of mathematical modelling for magnetic energy harvesting [2, 3], especially in the context of the Clampless Current-Transformer (CCT) method applied to induction motors. This study, therefore, embarks on a journey to bridge this gap by proposing a comprehensive mathematical modelling approach.

Previous research studies have delved into magnetic energy harvesting, yielding valuable insights into its potential applications. However, amidst this body of knowledge, a discernible gap pertains specifically to the mathematical modelling of magnetic energy harvesting from induction motors using the Clampless Current-Transformer (CCT) method [4]. Despite the strides made in this field, an evident void persists in comprehending how precisely this method can be harnessed to amplify the optimization of power harvesting from induction motors. With this backdrop in mind, our study assumes a crucial role as it seeks to transcend these limitations and contribute substantially to the current state of knowledge [5–9].

The research endeavors to bridge this gap by comprehensively exploring the mathematical modelling of magnetic energy harvesting from induction motors via the CCT method. While existing studies have ventured into the landscape of magnetic energy harvesting, they have predominantly focused on sources ranging from electrical transmission lines to power plants [10–12]. Intriguingly, the junction between the CCT method and induction motors has mainly remained uncharted, which this study meticulously designed to address it. The essence of this research gap serves as the driving force propelling the investigation forward, igniting the quest to unlock the latent potential of this method and decipher its intricate implications for optimizing power harvesting from induction motors [13–15]. Through this pursuit, this research aspires to illuminate unexplored dimensions, catalyzing a paradigm shift that underscores the pivotal role of mathematical modelling in realizing the maximal efficiency of energy utilization in this dynamic and transformative domain.

This article follows a structured approach to present the findings and analysis. Firstly, a comprehensive introduction establishes the research's background and significance. Then review the relevant literature, highlighting existing studies on magnetic energy harvesting and identifying the gaps our research aims to address. Next is the methodology employed, including the mathematical modelling approach and Ansys Engineering Simulation Software. The results of the simulations and

discussion of the implications for optimizing power harvesting from induction motors will be tabulated. Finally, the article summarizes the essential findings and their significance and suggests potential directions for future research in this field.

2 Materials and Method

This study employed specific materials for the core design, including Stainless Steel (SS) and Carbon Steel (CS), to evaluate their impact on power harvesting efficiency. Sensors, such as magnetic field sensors, were employed to measure and quantify the magnetic field strength emitted by the induction motor. The project has two significant designs: cylindrical and square core rods. Figure 1 shows all designs of cylindrical and square cores transducers with different numbers of turns. The specification of each design is tabulated in Table 1. All images in this study were taken from the experiment and software. Carbon steel is categorized as ferromagnetic low-carbon steel types. Most stainless-steel falls under the paramagnetic category, and this research used SS_304. In addition, CS and SS were chosen because they are readily available and low-cost.

The mathematical modeling and simulation were conducted systematically to ensure reliable and replicable results. The simulation was performed using Ansys Engineering Simulation Software, as shown in Fig. 2. The 3D outlines the sequential steps involved in the research, including the coil of core design, simulation setup, data collection, and analysis. The Steel_1008 represent the CS material of the core.

The simulation allows for precise modelling of electromagnetic transduction in the induction motor with the injected current according to the mathematical equation of the magnetic flux (Fig. 3).

The Eddy Current and Ultrasonic Tester were used to verify the CS and SS material used in the core design through the material's velocity and conductivity using the

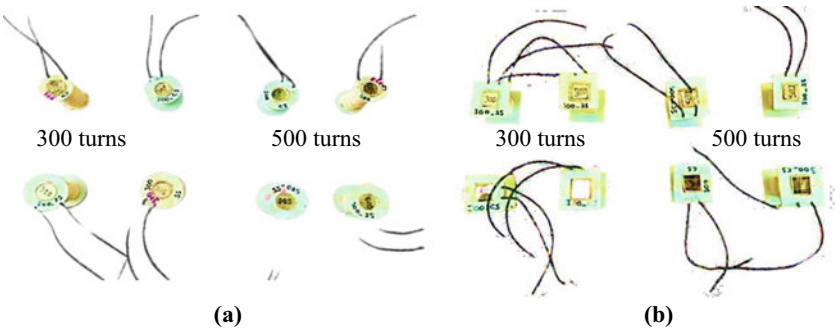


Fig. 1 Coil winding for SS and CS of different core shape **a** 300 turns and 500 turns cylindrical core, **b** 300 turns and 500 turns for square core

Table 1 The specification design of magnetic fields harvester for induction motor

No	Design of core	Material	No. of turns
1	Square short rod	Carbon steel	300
			500
2	Square long rod	Carbon steel	300
			500
3	Square short rod	Stainless steel	300
			500
4	Square long rod	Stainless steel	300
			500
5	Cylindrical short rod	Carbon steel	300
			500
6	Cylindrical long rod	Carbon steel	300
			500
7	Cylindrical short rod	Stainless steel	300
			500
8	Cylindrical long rod	Stainless steel	300
			500

Eddy Current tester. Figure 4 validated the material used through its ferromagnetic characteristics.

3 Result

The experimental result proves CS material is the best material for radiation and maximum current flow through the core. The low-frequency electromagnetic field makes hysteresis losses negligible, enabling the investigation of the magnetic properties of cores with different shapes. The experimental setup for these investigations is depicted in Fig. 5.

The Ansys Engineering Simulation result shows the radiation of flux density, B in unit Tesla and the relationship for magnetic flux with the different distances. Table 2 summarizes the voltage and current produced from each core design.

The highest output power is 500 cylindrical (long) turns for CS with 52.95 μW. The comparison of radiant flux density for cylindrical long between 300 and 500 turns is summarized in Fig. 6.

The 500 turns CS long rectangular core design produces lower output power than the cylindrical core with only 4.44 μW. Figure 7 compares the flux density rectangular core between 300 and 500 turns.

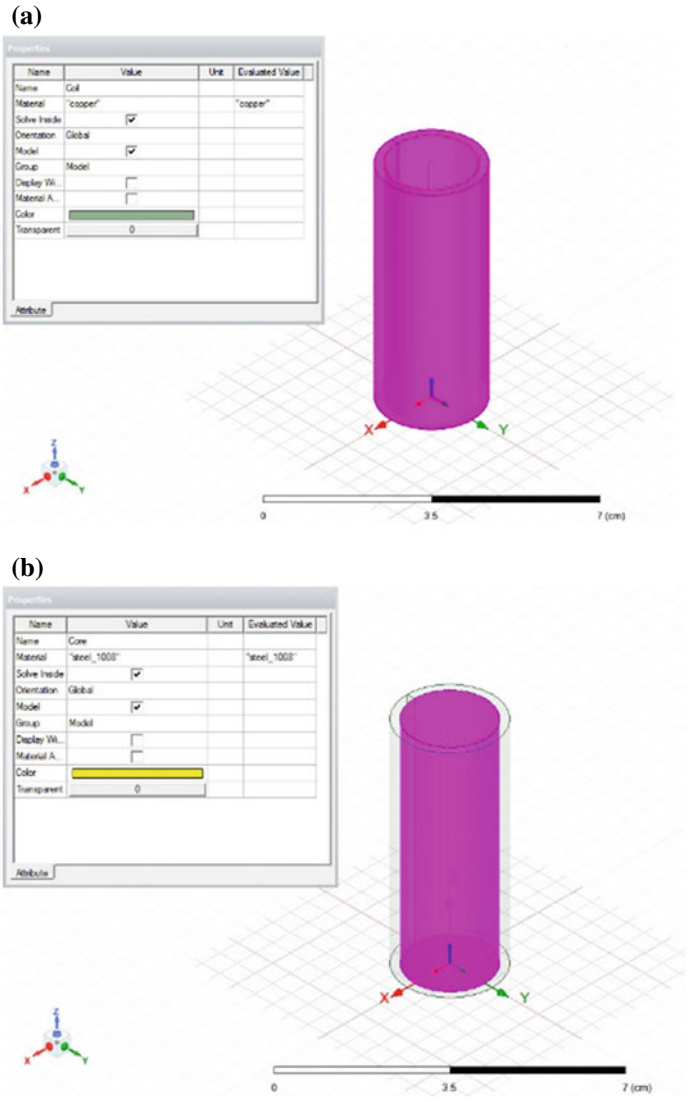


Fig. 2 The 3D of simulation **a** the coil of turning, **b** the core of design for cylindrical core shape

Radiant flux density refers to the amount of radiant energy passing through a unit area perpendicular to the direction of energy propagation. Figure 8 shows the flux density with 500 turns for cylindrical long with 0.0895 μT.

Compared with a long rectangular core, the flux density is only 0.0233 μT for 500 turns CS. The number of turns in a coil of wire can affect the flux density in specific situations. The relationship between the number of turns and the flux density

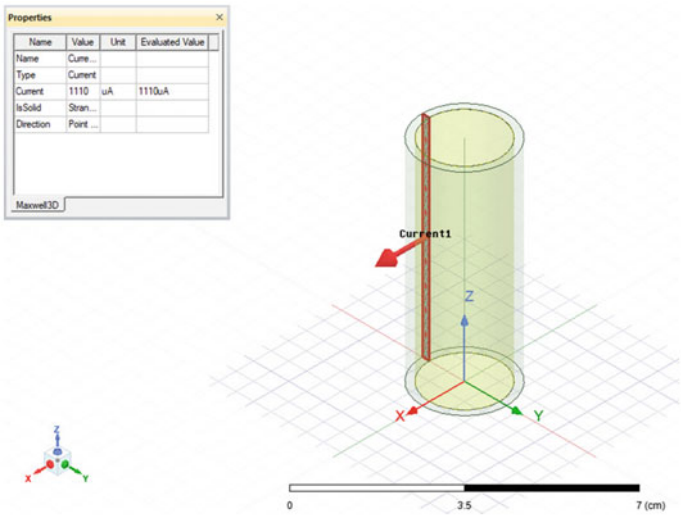


Fig. 3 The current direction of turning coil



Fig. 4 The Eddy current and ultrasonic setup. **a** The velocity of material using UT tester. **b** The conductivity of material using Eddy current tester

depends on the context of electromagnetic coils. Figure 9 shows the flux density with 500 turns for long rectangular CS.

Different materials exhibit varying levels of magnetic properties, such as permeability and saturation, influencing how they respond to magnetic fields and how efficiently they can be used in various applications. CS is the best magnetization under ferromagnetic characteristics in this simulation and experiment. Table 3 summarizes

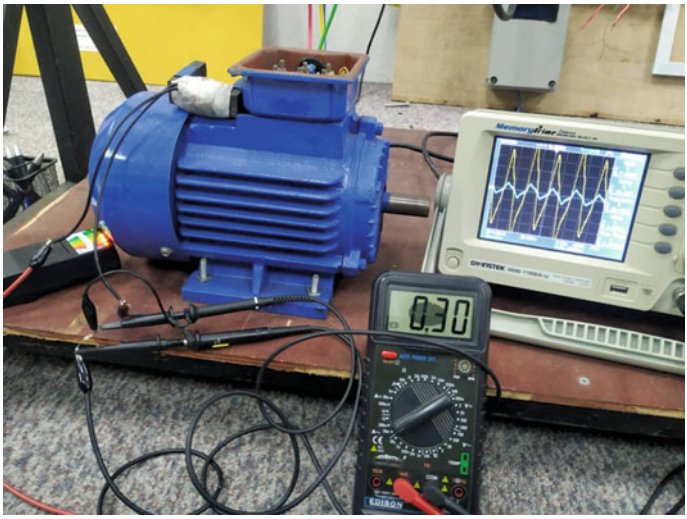


Fig. 5 Experimental setup

Table 2 The voltage and current of magnetic fields harvester for design of core

Material	Stainless Steel				Carbon Steel (Steel_1008)			
Turns	300		500		300		500	
Core Design	Voltage (mV)	Current (μA)	Voltage (mV)	Current (μA)	Voltage (mV)	Current (μA)	Voltage (mV)	Current (μA)
Cylindrical (short)	5.15	283	8.78	270	26.2	1080	30.0	1300
Cylindrical (long)	4.37	288	7.13	177	25.6	1200	47.7	1110
Rectangular (short)	3.91	236	8.35	532	17.4	598	23.7	975
Rectangular (long)	7.44	311	7.94	378	26.6	1230	22.0	1000

the difference in current injected in the coil, which makes a difference in magnetic flux in Tesla. The highest is the long core cylindrical with 500 turns producing 0.0895 μT.

Eddy currents are circulating currents induced in a conductive material when exposed to a changing magnetic field, typically alternating current. The behavior of a material's impedance in the frequency domain can give insights into how eddy currents interact with the material. Different materials will have different impedance responses depending on their electrical conductivity and other factors. Figure 10 shows CS, SS for the non-magnetic area and Aluminum for the magnetic area tested using Eddy Current Tester.

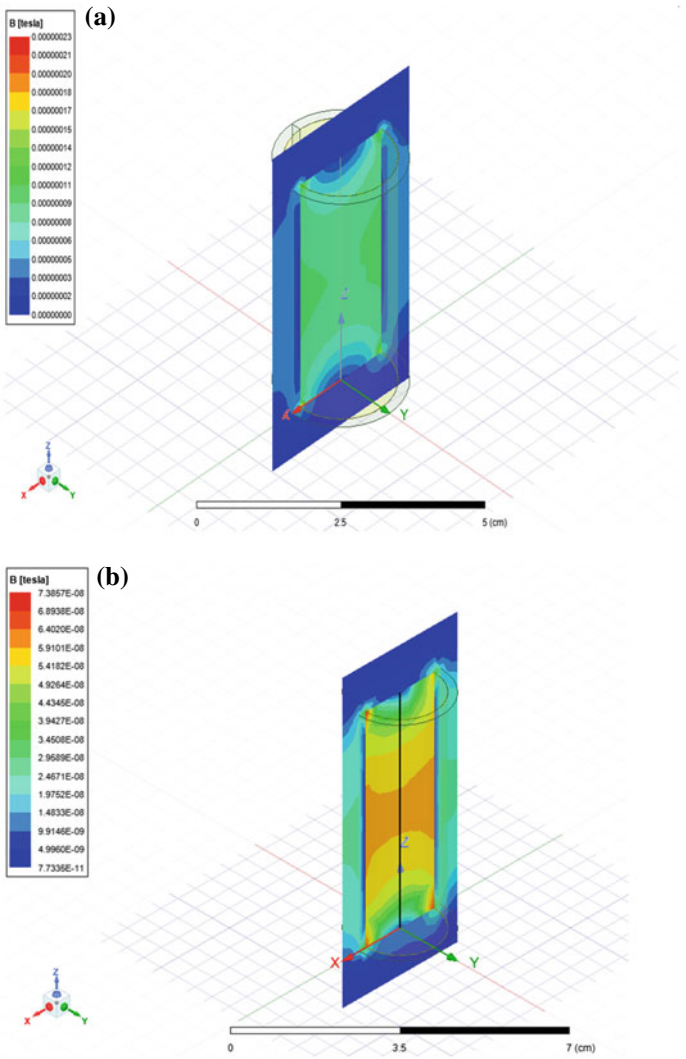


Fig. 6 Comparison of radiant flux density for cylindrical long core **a** with 300 turns, **b** with 500 turns

The electrical conductivity of material influences the magnitude and distribution of eddy currents. Conductive materials allow the eddy currents to flow more easily. In eddy current testing, the change in impedance of the coil as it interacts with the material's eddy current can be used to determine its conductivity. Figure 11 compares the conductivity of CS and SS, with CS producing better current than SS with 5.05%IACS.

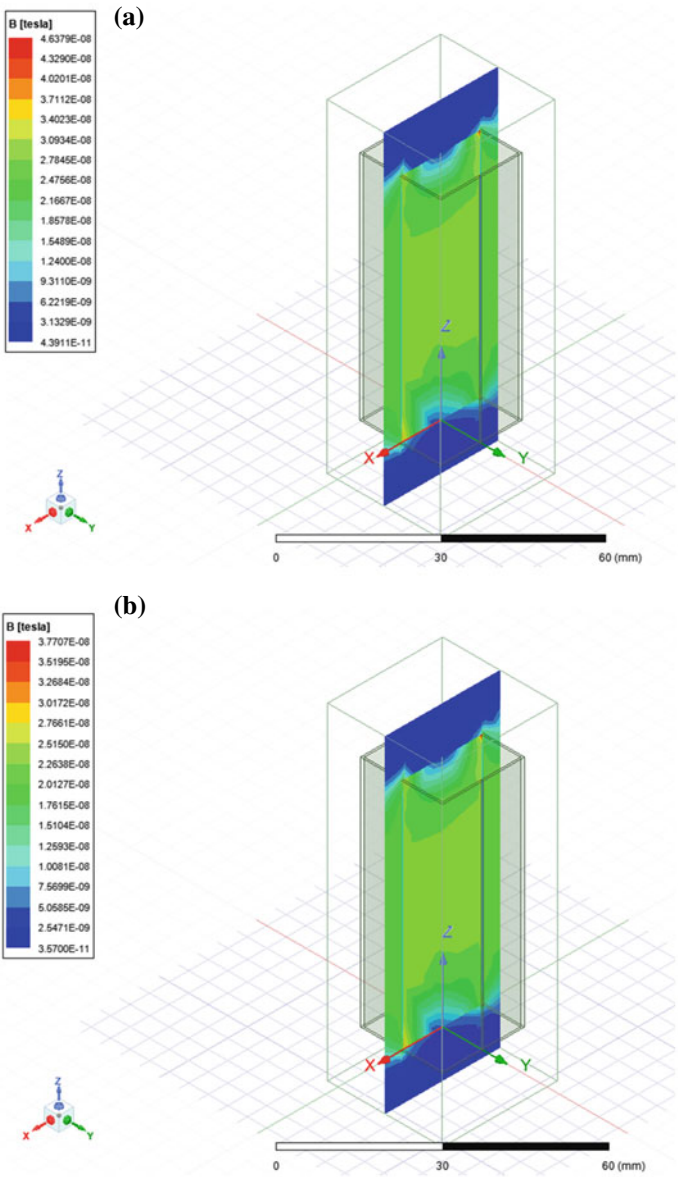


Fig. 7 The comparison of radiant flux density for long core rectangular **a** with 300 turns, **b** with 500 turns

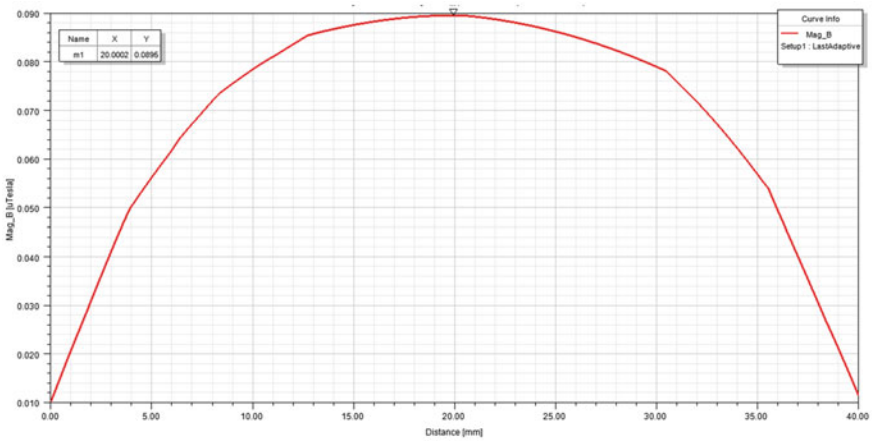


Fig. 8 Flux density for CS long cylindrical with 500 turns

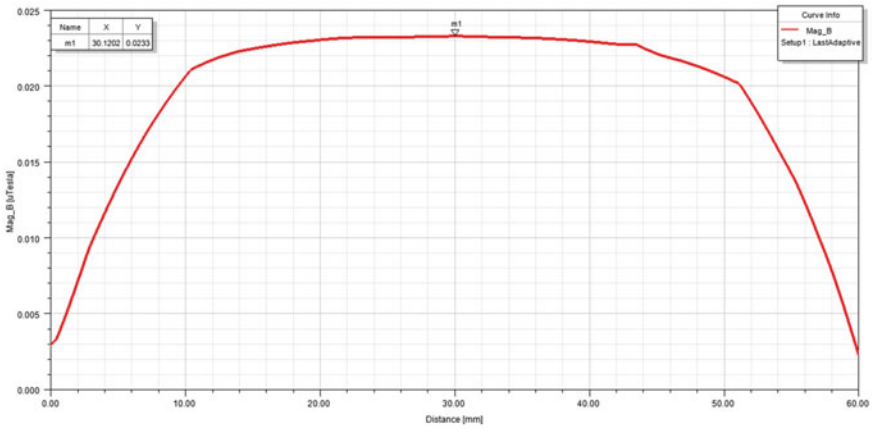


Fig. 9 Flux density for long rectangular CS with 500 turns

Ultrasonic testing (UT) involves sending sound waves with frequencies above the audible range into a material and measuring the time it takes to travel through it and reflect. A transducer usually generates these waves, travelling as longitudinal waves through the material. Determining velocity in UT is more commonly used for calibrating the equipment and obtaining accurate measurements of the material's thickness or detecting flaws based on the time the waves take to travel. Figure 12 shows the comparison of the velocity of CS and SS. The velocity of CS is 6012 m/s (SI for Mild Steel 5900 m/s), and for SS is 5797 m/s (SI for Stainless Steel 5800 m/s). Thus the reading shows that CS velocity is better compared to SS.

Table 3 The output power and flux density of magnetic fields harvester for design of core

Material	Carbon Steel (Steel_1008)					
Turns	300			500		
Core Design	Current (μA)	Power Output (μW)	Flux Density (μT)	Current (μA)	Power Output (μW)	Flux Density (μT)
Cylindrical (short)	1080	28.30	0.0596	1300	39.00	0.0550
Cylindrical (long)	1200	30.72	0.0745	1110	52.95	0.0895
Rectangular (short)	598	10.41	0.0209	975	23.11	0.0341
Rectangular (long)	1230	22.00	0.0286	1000	32.72	0.0233

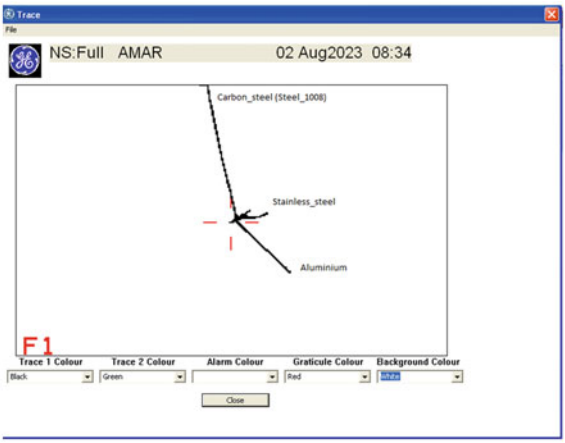


Fig. 10 The Eddy current impedance plane responses



(a) **(b)**

Fig. 11 Comparison of conductivity by using Eddy current tester **a** carbon steel, **b** stainless steel



(a) **(b)**

Fig. 12 Comparison of velocity using UT tester **a** CS, **b** SS

4 Discussion

The mathematical modelling indicates that different core designs significantly impact the magnetic flux density. For instance, 500 turns long cylindrical demonstrated the highest magnetic flux density with a 0.0895 μT, while 500 turns long rectangular exhibited the lowest flux density of 0.0233 μT. The observed differences highlight the importance of carefully considering and optimizing these factors to optimize magnetic waste and enhance the system's overall efficiency. Even if limited studies were performed on magnetic energy harvesting from manufacturing machinery [16–18], such as the induction motor, the electromagnetic transducer employed the clamped current-transformer technique [19–23].

Compared to other related studies, this research contributes to understanding magnetic waste intensity in material properties and complements the existing literature by providing additional insights into the specific parameters and conditions that influence magnetic waste intensity. Furthermore, using Eddy current and ultrasonic testing allows for more accurate and detailed analysis, enhancing the reliability of the findings.

Despite the valuable insights gained from this research, several limitations should be acknowledged. Firstly, the accuracy of the results heavily relies on the input parameters and assumptions used in the modelling process. Sensitivity analysis and parameter optimization could help mitigate these limitations and improve the reliability of the results. Finally, the study is limited to specific simulation scenarios and core designs. Including a broader range of scenarios and designs would enhance the comprehensiveness and applicability of the study.

This study holds several implications for magnetic flux with the core of the design. By providing insights into the factors influencing magnetic waste intensity, it offers guidance for designing and optimizing output power harvesting. The findings emphasize the importance of considering core design choices and simulation scenarios to minimize magnetic waste and improve overall system efficiency, which has implications for energy conservation, sustainability, and cost-effectiveness in various industries and applications reliant on magnetic systems. Moreover, the research

contributes to the broader understanding of magnetic waste reduction strategies, fostering advancements in technology and innovation to enhance energy efficiency and minimize environmental impact.

5 Conclusion

This study explored the mathematical modelling and simulation of output power and magnetic flux density, focusing on the impact of different core designs and simulation scenarios. The analysis revealed significant variations in magnetic waste intensity among the simulation scenarios. Future research should aim to conduct sensitivity analysis to identify the most influential parameters, and optimizing these parameters could further enhance the accuracy and reliability of the mathematical modelling. Exploring the impact of a broader range of simulation scenarios and core designs would provide a more comprehensive understanding of the magnetic waste intensity and guide the development of strategies to minimize waste. In summary, future research may address integrating IoT and intelligent systems, energy storage solutions, robustness and reliability and optimization algorithms, which can lead to more efficient, practical, and sustainable solutions, significantly impacting energy efficiency and conservation in various industries.

References

1. Cai M, Liao W-H (2021) Design, modeling, and experiments of electromagnetic energy harvester embedded in smart watch and wristband as power source. IEEE/ASME Trans Mechatron 26:2104–2114
2. Mencou S, Yakhlef MB, Tazi EB (2022) Advanced torque and speed control techniques for induction motor drives: a review. In: 2022 2nd international conference on innovative research in applied science, engineering and technology (IRASET). IEEE
3. Juez-Gil M et al (2020) Early and extremely early multi-label fault diagnosis in induction motors. ISA Trans 106:367–381
4. Riba J-R, Moreno-Eguilaz M, Bogarra S (2022) Energy harvesting methods for transmission lines: a comprehensive review. Appl Sci 12(21):10699
5. Kato M (2023) Numerical simulation on electromagnetic energy harvester oscillated by speed ripple of AC motors. Energies 16(2):940
6. Mounir M, Adil E, Nabil C, Mohamed I (2021) The use of magnetostrictive intelligent sensors for fault tolerant control of induction motors with energy harvesting principle. Trans Inst Meas Control 43(9):1996–2005
7. Haniszewski T, Cieśla M (2022) Energy harvesting in the crane-hoisting mechanism. Energies 15(24):9366
8. Curto D et al (2020) A new solution for sea wave energy harvesting, the proposal of an ironless linear generator. J Mar Sci Eng 8(2):93
9. Bijak J et al (2022) Modelling and investigation of energy harvesting system utilizing magnetically levitated permanent magnet. Sensors 22(17):6384
10. Putrada AG et al (2022) Machine learning methods in smart lighting toward achieving user comfort: a survey. IEEE Access 10:45137–45178

11. Kumar A et al (2021) 6th generation: communication, signal processing, advanced infrastructure, emerging technologies and challenges. In: 2021 6th international conference on computing, communication and security (ICCCS). IEEE
12. Saranyaa JS, Peer Fathima A (2023) A comprehensive survey on the current trends in improvising the renewable energy incorporated global power system market. IEEE Access
13. Peng X, Liu Z, Jiang D (2021) A review of multiphase energy conversion in wind power generation. Renew Sustain Energy Rev 147:111172
14. Wang J et al (2020) The state-of-the-art review on energy harvesting from flow-induced vibrations. Appl Energy 267:114902
15. Sarker MR et al (2019) Review of piezoelectric energy harvesting system and application of optimization techniques to enhance the performance of the harvesting system. Sens Actuators A Phys 300:111634
16. Sun Q, Patil S, Sun NX, Lehman B (2013) Modeling and optimization of an inductive magnetic harvester considering nonlinear effects. https://doi.org/10.1109/COMPEL.2013.6626463
17. Wright SW, Kiziroglou ME, Spasic S, Radosevic N, Yeatman EM (2019) Inductive energy harvesting from current-carrying structures. IEEE Sens Lett 3(6). https://doi.org/10.1109/LSENS.2019.2918339
18. Kiziroglou ME, Wright SW, Yeatman EM (2022) Power supply based on inductive harvesting from structural currents. IEEE Internet Things J 9(10):7166–7177. https://doi.org/10.1109/JIOT.2021.3086013
19. Tashiro K, Wakiwaka H, Inoue SI, Uchiyama Y (2011) Energy harvesting of magnetic powerline noise. IEEE Trans Magn 47(10):4441–4444. https://doi.org/10.1109/TMAG.2011.2158190
20. van Schalkwyk JA, Hancke GP (2012) Energy harvesting for wireless sensors from electromagnetic fields around overhead power lines. In: IEEE international symposium on industrial electronics, pp 1128–1135. https://doi.org/10.1109/ISIE.2012.6237247
21. dos Santos MP, Vieira DA, Rodriguez YPM, de Souza CP, de Moraes TO, Freire RCS (2014) Energy harvesting using magnetic induction considering different core materials. In: Conference record - IEEE instrumentation and measurement technology conference, pp 942 944. https://doi.org/10.1109/I2MTC.2014.6860881
22. Wang J, Kim J, Ha DS (2021) Powerline energy harvesting circuit with a desaturation controller for a magnetic core. In: Midwest symposium on circuits and systems, vol. 2021-August, pp 220–223. https://doi.org/10.1109/MWSCAS47672.2021.9531890
23. Sánchez PM, Sánchez FJR, Gómez ES (2020) An experimental strategy for characterizing inductive electromagnetic energy harvesters, 20(3). https://doi.org/10.3390/s20030647

Comparison of CNN and KNN Methods for Cataract Classification and Detection Based on Fundus Images



Farah Hanifah, Nur Alifia Azzahra, Yunendah Nur Fuadah, Alvian Pandapotan, Erni Yanthy, Rita Magdalena, and Sofia Saidah

Abstract Cataract is an ophthalmic condition where the lens of the eye becomes cloudy, leading to symptoms such as blurry vision, increased sensitivity to light, nearsightedness, and blindness. The World Health Organization (WHO) states that 40% of blindness cases in the world are caused by cataracts. The rate of blindness caused by cataracts in developed countries is 5% while for countries and/or remote areas it is as high as 50%. This leads to a decrease in the productivity level of the society. Therefore, early detection plays an important role in this regard as it can help patients recognize the cataract at an early stage and take action according to the level of cataract experienced. The purpose of this study is to create an automated model to detect and classify cataracts into four conditions, normal, immature, mature, and hypermature using machine learning algorithms. In order to obtain an optimal model to detecting and classifying cataract, this study compares the performance of two machine learning algorithms, Convolutional Neural Network with the proposed layer and K-Nearest Neighbor. In this study, the dataset used was 2000 fundus images which is divided into 1600 training data images and 400 test data images taken from Cicendo Hospital, Garut, West Java, Indonesia. In the previous study, the cataract classification and detection system were carried out using a 3-Layer Convolutional Neural Network by classifying two conditions, normal and cataract with an accuracy value of 95%. Model performance is showed by confusion matrix analysis which includes accuracy, precision, recall, and F1-Score. The best performance is obtained when using 5-Layer Convolutional Neural Network with 98% system accuracy with hyperparameter 100 of epoch, 64 of batch size, and 0.001 of learning rate. Meanwhile, the system accuracy obtained by the K-NN method is 97% with Euclidean Distance, k = 3 and 0° of angle orientation. In this study it can be concluded that the classification system and cataract detection through fundus images obtain good results with the Convolutional Neural Network algortima with an accuracy value of 98%.

F. Hanifah · N. A. Azzahra · Y. N. Fuadah (✉) · A. Pandapotan · E. Yanthy · R. Magdalena · S. Saidah
Telkom University, Bandung, Indonesia
e-mail: yunendah@telkomuniversity.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_9

Keywords Cataract · Classification · Fundus image · Convolutional neural network (CNN) · K-nearest neighbor (KNN)

1 Introduction

The World Health Organization (WHO) claims that cataracts affect the vision of people of suffer, including blindness. About 51% of blindness cases in the world are caused by cataracts [8, 16]. In developed countries, the blindness rate caused by cataracts is around 5%. Meanwhile, the blindness rate in rural areas is as high as 50% [8]. In clinical terms, cataract is a condition where the loss of crystalline transparency in the lens of the eye causes the protein in the lens to become clumped together [25]. This condition can lead to eye problems, such as blurred and hazy vision, sensitivity to light, and blindness [23]. The initial blurring of vision will gradually lead to clouding of the lens, leading to decreased vision [16]. Cataracts are divided into two types, which are congenital cataracts and cataracts caused by other factors. Congenital cataract is a type of cataract that occurs from the birth or is genetic. Cataracts caused by other factors can occur due to aging, trauma such as hard punches, exposure to high heat, and exposure to chemicals [24]. Cataracts can be differentiated based on the level of maturity, which are immature, mature, and hypermature [2, 23]. Immature cataract is the first level of cataract where the patient's condition cannot accommodate near object clearly, because some of the lens is cloudy [22]. Mature cataract is an advanced level after immature cataract where the lens of the eye is cloudy as a whole. Finally, hypermature cataract is an advanced stage of mature cataract where fluid oozes out of the lens and causes changes to the lens [13].

In an attempt to reduce and prevent the risk posed by cataracts, early detection for cataracts plays an important role. In addition, it helps cataract affected people to recognize their condition earlier and take appropriate action. Cataract examinations are generally performed by ophthalmologists using a device called a slit lamp [6]. However, the examination cannot be performed in all health facilities due to equipment limitations. Therefore, an automated system for cataract detection is a solution to perform early cataract detection where the system model is designed using machine learning algorithms.

Based on previous research titled Cataract Classification Based on Fundus Images Using Convolutional Neural Network using several architectures and obtained the best accuracy of 93% with ResNet architecture to classify cataracts into four conditions, normal, immature, mature, and hypermature using a total dataset of 3200 data in total [18]. Likewise, in a study titled An Optimized Convolutional Neural Network (CNN) using Diffgrad for Cataract Image Classification using AlexNet which get an accuracy value of 97.5% from a dataset of 200 fundus image data sourced from the Kaggle Website [19].

In previous research with the title Cataract Detection in retinal Fundus Image Using Grey Level Co-occurrence Matrix and K-Nearest Neighbor, the classification

of cataracts into two conditions, normal and cataracts using the GLCM method, the value of k = 5 and the calculation of the Euclidean distance obtained an accuracy of 80% [15].

This research will compare methods for the classification and detection system of cataract fundus images that will be categorized based on four conditions, normal, immature cataracts, mature cataracts, and hypermature cataracts. The methods that will be used are 5-Layer Convolutional Neural Network by making hyperparameters a reference to find optimal system performance and the K-Nearest Neighbor method with GLCM feature extraction with distance parameters and the value of neighborliness used as a reference to determine optimal performance.

2 Method

The designed cataract classification system has a workflow as shown in Fig. 1. This system is designed using machine learning with the Convolutional Neural Network method using a basic 5-layer architecture that aims to obtain an optimal model in classifying cataract eye conditions into four condition, normal, immature, mature, and hypermature cataracts. The first process begins with preprocessing the fundus image in the form of augmentation and resizing. Augmentation is one of the techniques commonly used in the preprocessing stage where the original data is varied to improved image quality by rotation, shifting, mirroring, and other processes. The augmentation stage aims to avoid overfitting the system model [17]. While resizing aims to equalize the size of the fundus image. After preprocessing, the dataset will be classified using the 5-Layer Convolutional Neural Network method which consists of 5-layer convolutions.

In addition, this research also uses the K-Nearest Neighbor method with the working principle of finding test data that has the closest distance to k neighbors in the training data [14]. The KNN system design block diagram is shown in Fig. 2.

This research will compare two machine learning methods, Convolutional Neural Network and K-Nearest Neighbor for cataract classification systems using the same dataset.

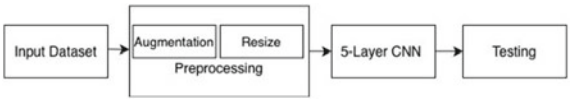


Fig. 1 System block diagram

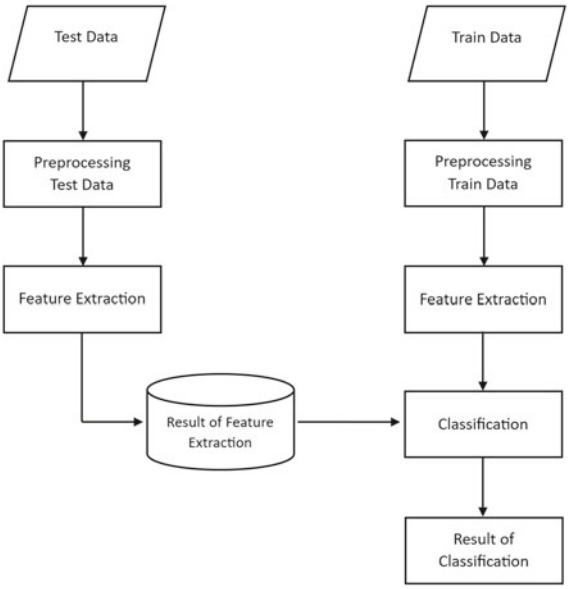
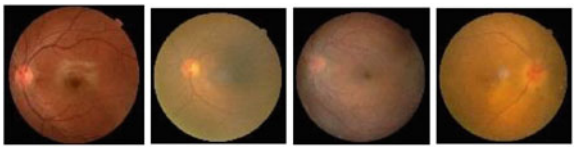


Fig. 2 K-nearest neighbor system block diagram

2.1 Dataset

This study uses a dataset in the form of fundus images of cataract eye conditions consisting of four conditions, normal, immature, mature, and hypermature cataracts. The dataset is obtained from Cicendo Garut Hospital with a total of 42 images of immature cataracts, 18 images of mature cataracts, and 9 images of hypermature cataracts. To create a cataract classification and detection model using machine learning algorithms, a huge amount of data is needed so that the machine learning model can train the data to recognize the type of cataract. Augmentation is one of the techniques commonly used in the preprocessing phase where the original data is varied to improve image quality by rotation, shifting, mirroring, and other processes. The augmentation process performed on the dataset consist of counterclockwise rotation by 180°, clockwise rotation by 180°, horizontal flip, vertical flip, noise addition, image blurring, and shift by (0, 40) where the image will be shifted by 40-pixels towards the y-axis and the action is repeated until the dataset totals 500 images for each class. The original datasets of the four classification classes of this study can be seen in Fig. 3.

The total dataset used for this classification system is 2000 cataract fundus image dataset, where for testing with both methods, namely Convolutional Neural Networks and K-Nearest Neighbor, 1600 images will be used for training data and 400 images for test data.



(a) (b) (c) (d)

Fig. 3 **a** Normal, **b** immature, **c** mature, **d** hypermature

2.2 Convolutional Neural Network

Convolutional Neural Network is one of the machine learning algorithms that excel in image processing in several fields, such as disease classification and traffic sign recognition [20]. This method has also developed in identifying, detecting, and classifying objects [7, 20]. Convolutional Neural Network is composed of three types of layers, convolutional, pooling, and fully connected layers [5].

Convolutional Layer

Convolutional layer is the main part of the convolutional neural network method which is the first layer. This layer is the most important part where the input image will go through the feature extraction process using a filter matrix or kernel. This convolutional process has a result in the form of a feature map that contains information from the image that has been convolved [9].

Pooling Layer

The pooling process, also known as subsampling, is the process of reducing the size of the convolved feature map and can reduce the computational burden. Two types of pooling commonly used in CNNs are max pooling and average pooling. Max pooling takes the maximum value from the image, while average pooling takes the average value from the image [4].

Fully Connected Layer

This layer is the last layer of the Convolutional Neural Network. In this section there is a transformation process in the dimensions of the image in order to be classified according to the information obtained previously [9].

In this research, the method tested on the cataract classification system is the CNN algorithm with a simple architecture consisting of five hidden layers. Each layer consists of a convolutional layer and a pooling layer with a kernel size in each layer of 8, 16, 32, 64, and 128 respectively with a filter size used 5×5 in each layer and a ReLu activation function. The last part has two fully connected layers with 10-unit neurons and ReLu activation of four conditions of cataract fundus image. The 5-layer cnn block diagram is shown in Fig. 4.

To evaluate the performance results of the designed system, this research uses the k-fold cross validation method. This method works by taking k random data samples and testing the performance of the machine learning model that has been made [21].

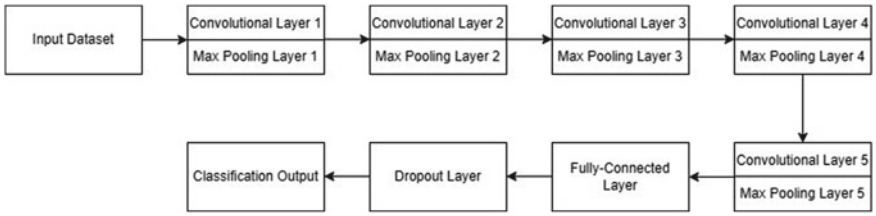


Fig. 4 5-layer convolutional neural network block diagram

This validation process is repeated k times with different data sample variations. The purpose of k-cross validation is to prevent overfitting [3].

2.3 K-Nearest Neighbor

K-Nearest Neighbor (K-NN) can be used to perform classification based on the number of K-Nearest Neighbor calculated using a distance function so that it can affect the accuracy results [11, 12]. The K value has an influence on the sensitivity of the machine learning model. The higher the value of k used, the noise effect can be reduced, but this can make the classification prediction boundaries more blurred and can result in a decrease in accuracy results [11, 21]. There are several distance functions including Euclidean, Minkowski, Chebychev, Correlation, Cityblock and Cosine.

To align the size of the fundus image that will be used in this research, resize with several sizes, such as 128×128, 256×256 and 512×512 pixels. Feature extraction in this study uses the Gray Level Co-occurance Matrix (GLCM) method which has a function to provide texture to the image in detail. Energy, Contrast, Correlation and Homogeneity are GLCM features used to classify mature cataracts, immature cataracts, hypermature cataracts and normal eyes with angular orientations of 0°, 45°, 90° and 135° and quantization levels of 8, 16 and 32.

2.4 System Performance

The performance of the designed classification system can be measured using a confusion matrix that contains accuracy, precision, recall, and F1-Score [10]. To get the system performance can use Eqs. 1, 2, 3, and 4.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{F1-Score} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \tag{4}$$

True Positive (TP) is a condition where the machine learning model can show positive data accurately predicted as a positive class. True Negative (TN) is a condition where the model can classify negative data accurately predicted as a negative class. Meanwhile, False Positive (FP) is a condition where the machine learning model misclassifies negative data as a positive class. This error is often referred to as type one error. False Negative (FN) can be referred to as type two error, where the model misclassifies positive data into a negative class [1].

3 Result

The cataract classification system in the process used 2000 cataract fundus image data from four conditions, namely normal, immature cataract, mature cataract, and hypermature cataract obtained from Cicendo Hospital Garut.

3.1 5-Layer Convolutional Neural Network

The system training process with the CNN method uses 2000 fundus images with divisions such as, 1600 images for training data, 400 images for test data, and 3200 images from training data used for validation data and for system validation a 5-fold cross validation method is used.

Several scenarios were formed in the training stage with the aim of determining the optimal model and knowing the effect of hyperparameters on system performance. The types of hyperparameters used for comparison are epoch, batch size, learning rate, and optimizer.

The first scenario is to determine the optimal epoch value for the system model. Epoch is the value of one round when the dataset has gone through the training process. The epoch values tested were 10, 25, 50, and 100. When conducting epoch tests, other hyperparameter values used are batch size of 64, learning rate of 0.001, and Adamax optimizer.

Table 1 shows the epoch effect on the performance of the CNN model in the cataract classification system. It can be seen, as the epoch value increases the performance of the system model is getting better. This proves that the epoch value affects

Table 1 Comparison of CNN model performance against epoch value

Epoch	Metric		Performance
10	Accuracy	0.7	0.6775
	Precision	0.66	
	Recall	0.7	
	F1-score	0.65	
25	Accuracy	0.91	0.91
	Precision	0.91	
	Recall	0.91	
	F1-score	0.91	
50	Accuracy	0.96	0.96
	Precision	0.96	
	Recall	0.96	
	F1-score	0.96	
100	Accuracy	0.98	0.98
	Precision	0.98	
	Recall	0.98	
	F1-score	0.98	

the performance of the machine learning model. When the epoch value is 10, the model performance obtained is 0.6775, this can occur because the model does not have enough iterations to achieve maximum performance. When the epoch value is 100, the model performance reaches 0.98 where this value is a better result than testing other epoch values. The increasing number of iterations helps the model train the data to achieve optimal results.

In the next scenario, the hyperparameter tested was batch size to determine the best batch size value for the system. The batch size values tested were 16, 32, and 64. During the testing process, the epich value used is 100, the learning rate is 0.001 and the optimizer is Adamax.

Batch size is the number of data samples taken from the training data to be tested at one time. Based on Table 2, the batch size value of 16 gives the best model performance result of 0.98.

The last test is to determine the optimal learning rate value for the system. The tested values are 0.1, 0.01, 0.001, and 0.0001. Other hyperparameters used in this test are epoch 100, batch size 64 and Adamax optimizer.

Table 3 presents the model performance values from the test results with several learning rate values. The best performance is given when the learning rate value is 0.001 learning rate is a measure of how fast the machine learning model performs data training. However, when the learning rate is 0.0001, the performance of the resulting model is smaller than the previous value. This can be caused by a large learning rate value that can exceed the optimal limit of a system so that the training process becomes unstable.

Table 2 Comparison of CNN model performance against batch size

Batch size	Metric		Performance
16	Accuracy	0.98	0.98
	Precision	0.98	
	Recall	0.98	
	F1-score	0.98	
32	Accuracy	0.97	0.9725
	Precision	0.98	
	Recall	0.97	
	F1-score	0.97	
64	Accuracy	0.97	0.97
	Precision	0.97	
	Recall	0.97	
	F1-score	0.97	

Table 3 Comparison of CNN model performance against learning rate

Learning rate	Metric		Performance
0.1	Accuracy	0.24	0.1575
	Precision	0.06	
	Recall	0.24	
	F1-score	0.09	
0.01	Accuracy	0.97	0.97
	Precision	0.97	
	Recall	0.97	
	F1-score	0.97	
0.001	Accuracy	0.98	0.98
	Precision	0.98	
	Recall	0.98	
	F1-score	0.98	
0.0001	Accuracy	0.74	0.74
	Precision	0.74	
	Recall	0.74	
	F1-score	0.74	

Testing the cataract fundus image classification system model by changing several hyperparameters, such as epoch value, batch size, and learning rate to get optimal system performance. After passing the testing stage, the optimal system performance was found when the epoch value was 100, batch size 64, learning rate 0.001, and Adamax optimizer. The results can be seen in Figs. 5 and 6 which states that the system model does not have underfitting and overfitting.

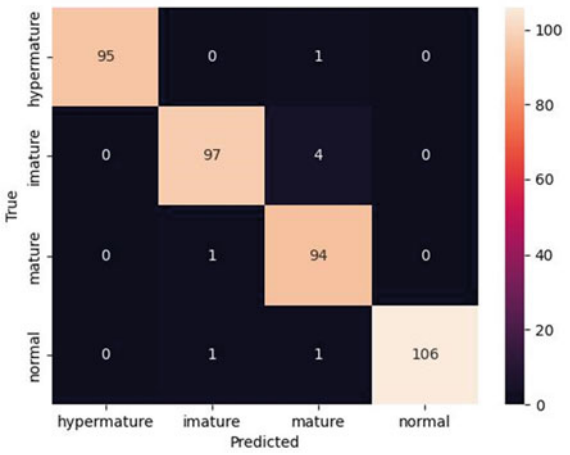


Fig. 5 CNN confusion matrix

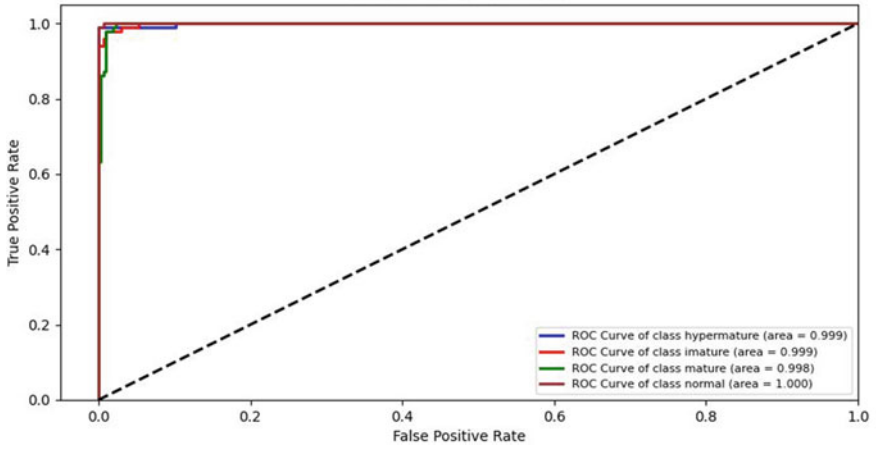


Fig. 6 ROC curve

3.2 K-Nearest Neighbor

In K-NN, the dataset used amounted to 2000 fundus images with a division of 1600 training data and 400 test data. The distance functions to be used in this comparison are Euclidean and Cosine with 256 × 256 resize, 90° angle orientation and 8 quantization levels.

Based on Table 4 in the comparison of performance calculations, it is stated that the results are not significantly different from each experiment. However, K-NN using the Euclidean distance function with the number of neighbors 3 produces the best performance of 0.9765.

Table 4 Performance comparison of K-NN model

Distance	NN	Metric		Performance
Euclidean	3	Accuracy	0.9775	0.9765
		Precision	0.975	
		Recall	0.9775	
		F1-score	0.9762	
	5	Accuracy	0.96	0.959
		Precision	0.9575	
		Recall	0.96	
		F1-score	0.9587	
	7	Accuracy	0.9475	0.9465
		Precision	0.945	
		Recall	0.9475	
		F1-score	0.9462	
	9	Accuracy	0.95	0.949
		Precision	0.9475	
		Recall	0.95	
		F1-score	0.9487	
Cosine	3	Accuracy	0.975	0.9731
		Precision	0.97	
		Recall	0.975	
		F1-score	0.9724	
	5	Accuracy	0.955	0.954
		Precision	0.9525	
		Recall	0.955	
		F1-score	0.9537	
	7	Accuracy	0.9425	0.94
		Precision	0.9375	
		Recall	0.9425	
		F1-score	0.9399	
	9	Accuracy	0.9325	0.9296
		Precision	0.925	
		Recall	0.9325	
		F1-score	0.9287	

Table 5 Performance comparison K-NN and CNN methods

Algorithm	Metric		Performance
KNN 3 Euclidean	Accuracy	0.9775	0.9765
	Precision	0.975	
	Recall	0.9775	
	F1-score	0.9762	
5-Layer CNN	Accuracy	0.98	0.98
	Precision	0.98	
	Recall	0.98	
	F1-score	0.98	

4 Discussion

Table 5 shows the performance results of the cataract fundus image classification system show that the CNN algorithm produces more optimal performance than the K-NN algorithm with a performance value of 0.98. In CNN testing, optimal results are obtained when using hyperparameters, namely epoch value 100, batch size 64, learning rate 0.001 and adamax optimizer. Meanwhile, the system performance generated in testing with the K-NN algorithm is 0.9765 with parameters with the number of neighbors 3 and the Euclidean distance function.

The performance results produced by the two algorithms have a small difference, based on the results listed CNN algorithm is superior for having higher F-1 Score results compared to the K-NN algorithm in classifying cataract fundus images, as evidenced by the results of Accuracy, Precision and Recall.

5 Conclusion

In the trial of classifying normal eye fundus images, immature cataracts, mature cataracts and hypermature cataracts using the K-Nearest Neighbor (K-NN) and Convolutional Neural Network (CNN) algorithms, it was found that the 5-Layer Convolutional Neural Network method produces optimal performance at a value of 0.98 or 98% of the accuracy, precision, recall and f1-Score values. The performance results are affected by the hyperparameter values used as a reference, such as epoch, batch size, and learning rate. Optimal system performance is obtained when the epoch value is 100, batch size 64, and learning rate 0.001. Meanwhile, the K-NN test resulted in a performance of 0.9765 or 97.65%. The parameters used in the test use Gray Co-occurance Level Matrix (GLCM) feature extraction with 90° angle orientation, 8 levels of quantization and resize by 256 × 256 and use the Euclidean distance function with the number of neighbors as many as 3. After comparing between CNN and KNN, the maximum performance is produced by CNN in classifying and

detecting cataracts. For future system development, comparisons can be made with other machine learning models to get better optimal results.

References

1. Anggreany DMS (2020) Confusion matrix. https://socs.binus.ac.id/2020/11/01/confusion-matrix/
2. Belikov AV et al (2020). A classification: using back propagation neural network algorithm to identify cataract disease a classification: using back propagation neural network algorithm to identify cataract disease. https://doi.org/10.1088/1742-6596/1566/1/012037
3. Berrar D (2018) Cross-validation. Encycl Bioinforma Comput Biol ABC Bioinforma 1–3, January 2018, 542–545. https://doi.org/10.1016/B978-0-12-809633-8.20349-X
4. Bezdan T, Bacanin N (2019) Convolutional neural network layers and architectures. https://doi.org/10.15308/Sinteza-2019-445-451
5. Bhatt D et al (2021) CNN variants for computer vision: history, architecture, application, challenges and future scope. Electronics 10(20):1–28. https://doi.org/10.3390/electronics10202470
6. Cao L et al (2020) Hierarchical method for cataract grading based on retinal images using improved Haar wavelet. Inf Fusion 53:196–208. https://doi.org/10.1016/j.inffus.2019.06.022
7. Chauhan R (2018) Convolutional neural network (CNN) for image detection and recognition. In: 2018 first international conference on secure cyber computing and communication, pp 278–282. https://doi.org/10.1109/ICSCCC.2018.8703316
8. Chen X et al (2021) Cataract: advances in surgery and whether surgery remains the only treatment in future. Adv Ophthalmol Pract Res 1(1):100008. https://doi.org/10.1016/j.aopr.2021.100008
9. Dhillon A, Verma GK (2020) Convolutional neural network: a review of models, methodologies and applications to object detection. Prog Artif Intell 9(2):85–112. https://doi.org/10.1007/s13748-019-00203-0
10. Düntsch I, Gediga G (2019) Confusion matrices and rough set data analysis. J Phys Conf Ser 1229:1. https://doi.org/10.1088/1742-6596/1229/1/012055
11. Gou J et al (2019) A generalized mean distance-based K-nearest neighbor classifier. Expert Syst Appl 115:356–372. https://doi.org/10.1016/j.eswa.2018.08.021
12. Hatuwal B et al (2021) Plant leaf disease recognition using random. https://doi.org/10.17562/PB-62-2
13. Kulkarni C, Metri RM (2019) Comparison of nuclear size in mature and hypermature cataract. 5(4):426–431
14. Lubis AR, Lubis M (2020) Optimization of distance formula in K-nearest neighbor method. 9(1):326–338. https://doi.org/10.11591/eei.v9i1.1464
15. Miranti MG et al (2021) Cataract detection in retinal fundus image using gray level co-occurrence matrix and K-nearest neighbor. IJCSE (Atlantis Press) 209:268–271
16. Moghadam SH, Ganji J (2019) Evaluation of the nursing process utilization in a teaching hospital, Ogun State, Nigeria. J Nurs Midwifery Sci 6(3):149–155. https://doi.org/10.4103/JNMS.JNMS
17. Poojary R et al (2021) Effect of data-augmentation on fine-tuned CNN model performance. IAES Int J Artif Intell 10(1):84–92. https://doi.org/10.11591/ijai.v10.i1.pp84-92
18. Simanjuntak RBJ et al (2022) Cataract classification based on fundus images using convolutional neural network. Int J Inform Vis 6(1):33–38. https://doi.org/10.30630/joiv.6.1.856
19. Sudarsono E et al (2020) An optimized convolutional neural network using diffGrad for cataract image classification. AIP Conf Proc 2296. https://doi.org/10.1063/5.0030746
20. Sun Y et al (2020) Evolving deep convolutional neural networks for image classification. IEEE Trans Evol Comput 24(2):394–407. https://doi.org/10.1109/TEVC.2019.2916183

21. Taunk K (2019) A brief review of nearest neighbor algorithm for learning and classification. In: 2019 International conference on intelligent computing and control systems ICCS 2019, ICICCS, pp 1255–1260
22. Uyanege PP, Abegunasekara NS (2021) Leech application in immature cataract
23. Xiong Y et al (2018) Automatic cataract classification based on multi-feature fusion and SVM. In: 2018 IEEE 4th International conference on computer and communications ICCC 2018, pp 1557–1561. https://doi.org/10.1109/CompComm.2018.8780617
24. Xu J et al (2020) Advances in pharmacotherapy of cataracts. Ann Transl Med 8(22):1552–1552. https://doi.org/10.21037/atm-20-1960
25. Zhang X et al (2022) Machine learning for cataract classification grading on ophthalmic imaging modalities: a survey. 19(June):184–208. https://doi.org/10.1007/s11633-022-1329-0

Design of Photoplethysmography (PPG)-Based Respiratory Rate Measuring Device Through Peak Calculations



Ummul Muthmainnah, Willy Anugrah Cahyadi, Husneni Mukhtar, Muhammad Abdul Hakiim Al Fatih, and Denny Tri Sukmono

Abstract Respiratory rate is one of the essential components in measuring vital signs that convey information about health conditions and diagnoses to patients measured in breaths per minute (BrPM). Breaths per minute refer to the number of breaths taken by a person in one minute. The current clinical respiratory rate calculation method, i.e., a standard calculation by counting the number of chest movements, can cause patient discomfort, bias in counting, and waste the nurse's time. This study aims to develop a tool for measuring respiratory rate by recording a one-minute photoplethysmography (PPG) signal non-invasively that is more comfortable and easier to use. The PPG signal obtained is first processed to remove unnecessary signal frequencies. The developed algorithm counts the number of peaks in the PPG signal to calculate the respiratory rate. It employs Zero-Phase Filtering to remove noise and motion artifacts (MA) on the signal to calculate the number of peaks in one minute. The device is designed using the MAX30102 sensor, XIAO ESP32-C3 controller, and TFT OLED display to show the measurement value. This tool uses a 3D case and Velcro tape on the sensor to reduce the possibility of a Motion Artifact appearing in the signal. The respiratory rate prediction test results obtained more than 94% accuracy with a maximum error of 3 BrPM in five subjects aged 15–48 years who were in good health. Based on the results, measuring respiratory rate from the PPG signal with a wearable device design can be a quite effective non-invasive solution and contribute to the medical world for measuring vital signs.

Keywords Respiratory rate · Photoplethysmography · Peak calculation

U. Muthmainnah · W. A. Cahyadi (✉) · H. Mukhtar · M. A. H. Al Fatih · D. T. Sukmono
School of Electrical Engineering, Telkom University, Bandung, Indonesia
e-mail: waczze@telkomuniversity.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_10

1 Introduction

Respiratory rate (RR) is critical to human vital sign monitoring. It refers to the number of RR a person takes in a minute and is a crucial indicator of their respiratory function [1, 2]. Healthcare providers monitor respiratory rate to detect respiratory problems and potential complications before they become serious. Additionally, changes in respiratory rate can indicate changes in a patient's health status, making it a critical measurement to monitor regularly [3]. The measurements can be obtained through invasive (harming the body) or non-invasive (not harming the body) techniques [4, 5]. However, respiratory rate can vary depending on age, gender, physical activity, and overall health. For example, newborn infants have a much higher respiratory rate than adults, while elderly individuals may have a slower respiratory rate due to age-related changes in lung function [6]. Monitoring respiratory rate is particularly important in patients with respiratory conditions such as asthma, chronic obstructive pulmonary disease (COPD), and pneumonia. In these patients, changes in respiratory rate can indicate worsening of their condition, and prompt medical intervention may be needed [7, 8].

Several studies have been conducted on devices for measuring respiratory rate from PPG signals in normal conditions, employing various methodologies. It employs PPG to analyze changes in blood volume in arteries near the skin of the fingers, enabling the calculation of respiratory rate [9, 10]. In 2014, Ungureanu C. and their team researched the respiratory rate extraction from PPG signals using a wearable device [11]. Their method involved observing PPG signal features, including PWV (Pulse Width Variability) or PAV (Pulse Amplitude Variability). Their developed way identified a signal feature, F11, which outperformed the Upper Pulse Envelope (F1) as a reference signal. To obtain the reference respiratory rate signal, they utilized data from Capnobase. They developed a respiratory rate device with a thermistor placed on a nasal cannula. The research yielded a maximum error of approximately 2 breaths per minute (BrPM) within a respiratory rate range of 6–30 BrPM. However, the instrument employed in this study involves a nasal cannula that may induce discomfort and serves merely as a measuring tool without displaying measurement results.

Subsequently, in 2018, Orlando S. and their team designed a Smartwatch for monitoring respiratory and heart rates using a combination of photosensors and a bandpass filter. They then extracted frequencies to determine the heart rate and respiratory rate. The wearable device they created for measuring respiratory rate exhibited a measurement error of only 1 BrPM [12]. The developed device uses a custom circuit design to acquire and process PPG signals through various circuitry components, such as filters and amplifiers integrated into a printed circuit board.

There has been a growing advancement in methodologies aimed at extracting respiratory rate from photoplethysmography (PPG) signals in recent years. Notably, in 2023, a substantial body of research has leveraged the capabilities of Machine Learning (ML) techniques for predicting respiratory rates. Among these endeavors, a notable study conducted an exhaustive assessment of several algorithmic approaches,

including Support Vector Regression (SVR), Gaussian Process Regression (GPR), Ensemble Trees, and Decision Tree models. The primary objective was to determine the optimal algorithm in terms of RMSE and MAE for predicting respiratory rate values. The results of this investigation prominently highlighted the superiority of the Gaussian Process Regression technique, yielding remarkable RMSE and MAE values of up to 0.98 and 0.57, respectively [13].

In a separate study within the same domain, a distinct approach was undertaken involving the utilization of the Whale Optimization Algorithm in conjunction with a hybrid relational vector machine (WOA-HRVM). This composite methodology was applied to a dedicated training dataset, with the resultant model demonstrating the capability to predict respiratory rates. Noteworthy outcomes of this study encompassed the successful development of a predictive model for respiratory rates, yielding MAE and RMSE metrics of 0.71 and 0.99 BrPM on the training dataset and 1.24 and 1.79 BrPM on the testing dataset [14].

The research that employs machine learning often uses publicly available PPG signal datasets, like BIDMC and MIMIC [13–15]. This means there's a lack of focus on creating new tools to measure respiratory rate. However, developing these measurement tools is essential [16, 17]. They ensure accurate and consistent data monitoring in biomedical applications. Moreover, they can significantly improve the efficiency of medical research and clinical trials [18]. This study focuses on developing a wearable device with a wristwatch-like design to measure respiratory rate accurately. The device employs the MAX30102 sensor to capture the PPG signal and applies a Butterworth Bandpass filter and peak calculation methods. The objective is to ascertain the frequency of respiratory rate. By presenting measurement outcomes on the device's screen, this research aims to contribute significantly to the medical domain, facilitating the assessment of vital signs, precisely the respiratory rate.

2 Materials and Method

2.1 System Design

In this study, the device combines the XIAO ESP32-C3 powered by a 3.7 V Lithium Polymer Battery. Additionally, the device incorporates a MAX30102 sensor for signal detection and a 1.3-inch TFT OLED screen for display purposes. The MAX30102 sensor is an optical sensor that determines the cardiovascular system's condition by measuring changes in blood volume in skin tissue or a Photoplethysmogram (PPG) signal [9]. In its application, this signal utilizes a transmitter in infrared light and a photodetector located close together (reflection method) [19]. The sensor output signal will be sent to the microcontroller as a data acquisition and signal processing device. The outcome of the signal processing, represented as the count of respiratory rate, will be showcased on the TFT OLED screen. Figure 1 shows the block diagram where the PPG signal is obtained from the subject's fingertip.

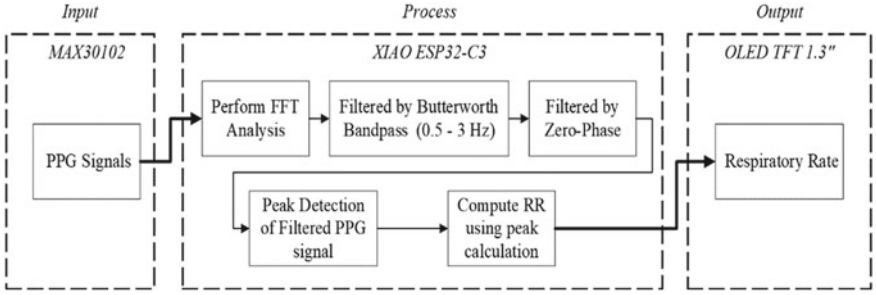


Fig. 1 PPG-based respiratory rate measurement system block diagram

The study thoroughly examined respiratory rates by analyzing the PPG signal. To identify the appropriate signal range for the Butterworth bandpass filter, the researchers conducted a Fast Fourier Transform (FFT) analysis. It could calculate the respiratory rate by counting the number of peaks within a minute in the filtered PPG signal. This analytical approach allowed for precisely determining the respiratory activity within a minute. The MAX30102 sensor was utilized to collect the PPG signal for this investigation.

2.2 Data Recording

This research employed PPG signals obtained from the recording output sensor for one minute to ensure accurate respiratory rate prediction results. The data was collected from five subjects with diverse physiological or pathological conditions. A sampling frequency of 25 Hz was utilized to capture the PPG data. The subjects aged 15–48 and in a healthy state assumed a relaxed position and positioned their finger correctly on the sensor casing during the recording process, as shown in Fig. 7c.

To assess the distribution of data and the level of accuracy of a measuring device, each subject takes measurements ten times. This repeated measurement approach allows for a more comprehensive analysis of the data. By collecting multiple measurements, we can calculate various statistical measures, including the standard deviation, to provide insights into the variability and precision of the device's readings.

As a comparison with the predicted respiratory rate results, researchers will manually calculate the respiratory rate of the subjects when collecting data simultaneously. In calculating the respiratory rate, the researcher will see how often the subject's chest or abdomen rises in one minute using a clinically standard method [20].

2.3 Pre-processing

The collected data was analyzed using the Fast Fourier Transform (FFT), a mathematical algorithm to convert a signal from the time domain to the frequency domain [21]. This transformation allows us to observe the frequency components present in the PPG waveform. By utilizing FFT, we can gain insights into the frequency range of the PPG signal. To obtain the FFT of the PPG signal, the MATLAB function tool was employed, utilizing the following formula:

$$A(f_k) = \frac{1}{N} \sum_{n=0}^{N-1} a(t_n) e^{-i \frac{2\pi kn}{N}} \tag{1}$$

where t_n is discrete-time samples, f_k is discrete frequencies samples, N is the number of samples in an FFT time block, and the multiplication $\frac{1}{N}$ is the normalization factor.

The signal spectrum shows that the PPG signal frequency range is 0–10 Hz (see Fig. 2). To detect and calculate peak and reduce any background noise, the signal was denoised by using a third-order Butterworth Band-pass filter with passband 0.5–3 Hz to separate the heart rate signal [12, 22]. The detected peak represents the systolic peak when the volume of blood pumped by the heart reaches its highest point [23]. The choice of using a third-order Butterworth filter is influenced by the constraints and capabilities of microcontrollers, which are limited to performing computations up to the highest order of 3 for Butterworth filters. The cut-off frequency is determined by the AC component's frequency in the PPG signal linked to the pulse, typically exceeding 0.5 Hz (equivalent to 30 beats per minute or 1 Hz as 60 BPM) [24].

Zero-phase filtering is applied before performing peak signal calculation on the signal that has been band-pass filtered from 0.5 to 3 Hz. Zero-phase filtering is used to minimize the signal's motion artifacts and phase delay [25, 26]. The Eq. (2) describes the Zero-Phase Filtering formula.

$$\begin{aligned} y[n] = \frac{1}{a_0}(b_0 x[n] + b_1 x[n-1] + \cdots + b_k x[n-k] \\ - a_1 y[n-1] - a_2 y[n-2] - \cdots - a_j y[n-j] \end{aligned} \tag{2}$$

where $x[n]$ represents the input signal, $y[n]$ represents the output signal, k and j are the feed forward-feed backward filter orders, b_i represents the feedforward filter coefficients, and a_i represents the feed backward filter coefficients.

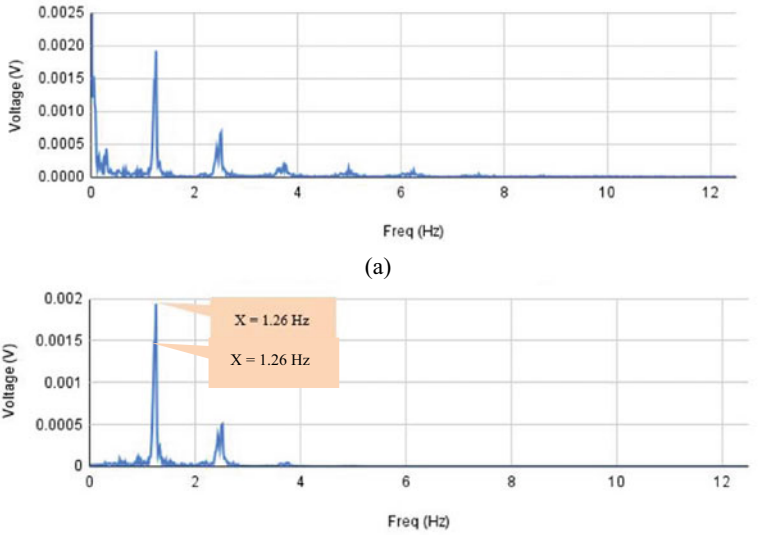


Fig. 2 **a** Frequency range of the original signal before filtering, **b** signal frequency after 0.5–3 Hz cut-off shows the frequency is 1.26 Hz, equal to 77 BPM

2.4 Respiratory Rate Extraction

The respiratory rate of a healthy adult (>20 years old) at rest usually ranges from 12 to 20 BrPM [3, 27]. Whereas at a younger age, such as teenagers aged 14–18 have a respiratory rate of 12–16 BrPM [20]. Respiration's primary role is to guarantee an ample supply of oxygen to the body's tissues and cells, supporting the process of aerobic metabolism while also eliminating carbon dioxide, the byproduct of metabolism [28].

The respiratory rate can be calculated from the number of peaks by eliminating unwanted peaks below a threshold value [29]. The number of peaks detected can also be interpreted as the number of pulse beats in one minute [30]. In this work, the PPG signal is automatically detected through a program in MATLAB. A threshold value of zero is established to exclude negative peaks from being considered valid peaks, thereby preventing the detection of false peaks [21, 31]. Figure 3 shows the detected peaks in one minute PPG signal.

A function is used to compare the values between t_n and t_{n-1} for the implementation of peak calculation on the microcontroller. Each value generated within one signal cycle is compared to detect the highest value. Each caught highest value corresponds to an increase in the peak count, which is then converted into the respiratory rate value.

$$RR = \frac{peaks}{4.5} \times t_n \tag{3}$$

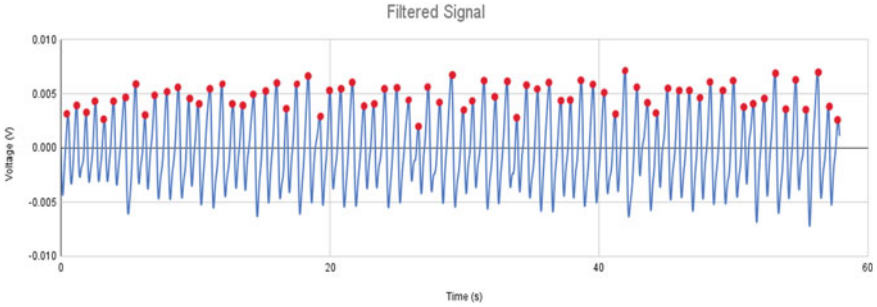


Fig. 3 One-minute filtered PPG signal with detected peak signatures, which represents a heart rate of 73 BPM

where RR is the value of respiratory rate, x is the number of peaks detected in one minute, and t_n is time normalization for one minute. Time normalization is fine-tuning the number of signal samples in one minute. Equation (3) shows the number of peaks divided by four based on the reference respiratory signal (flow) from the CapnoBase.org database. The flow compared to the PPG signal generated by the sensor showed that one breathing cycle contains 4.5 PPG signal cycles (hills and valleys) [11].

3 Result

The outcome of the tool development is a wearable watch featuring an enclosed 3D case design that enhances its aesthetic appeal Fig. 4. The device design considers the efficient placement of components. It incorporates various supporting features, including navigation buttons, a power switch, a battery compartment, and a micro-USB port.

In this study, the sensor is very vulnerable to movement. The slightest movement false sensor attachment can generate noise in the signal, called a motion artifact [24, 32]. Furthermore, motion artifacts can occur when the subject engages in speech activities. Movements of the tongue, lips, and jaw during speech can generate vibrations and slight fluctuations in blood flow [25, 33]. These fluctuations can disrupt the precise capture of the PPG signal, leading to inaccurate analysis and interpretation of the physiological information contained within the signal.

Hence, a 3D case was created for the sensor with a design adapted to place the fingertip. The creation of this 3D case aims to reduce finger pressure on the PPG sensor, which can reduce the quality of the PPG signal obtained [34]. Moreover, Velcro tape was also added to adhere the finger to the 3D case further, and it can be adjusted according to the size of the subject's finger, as shown in Fig. 5. It was evident that the signal generated without the 3D case and Velcro tape exhibited noticeable motion artifacts as shown in Fig. 6.

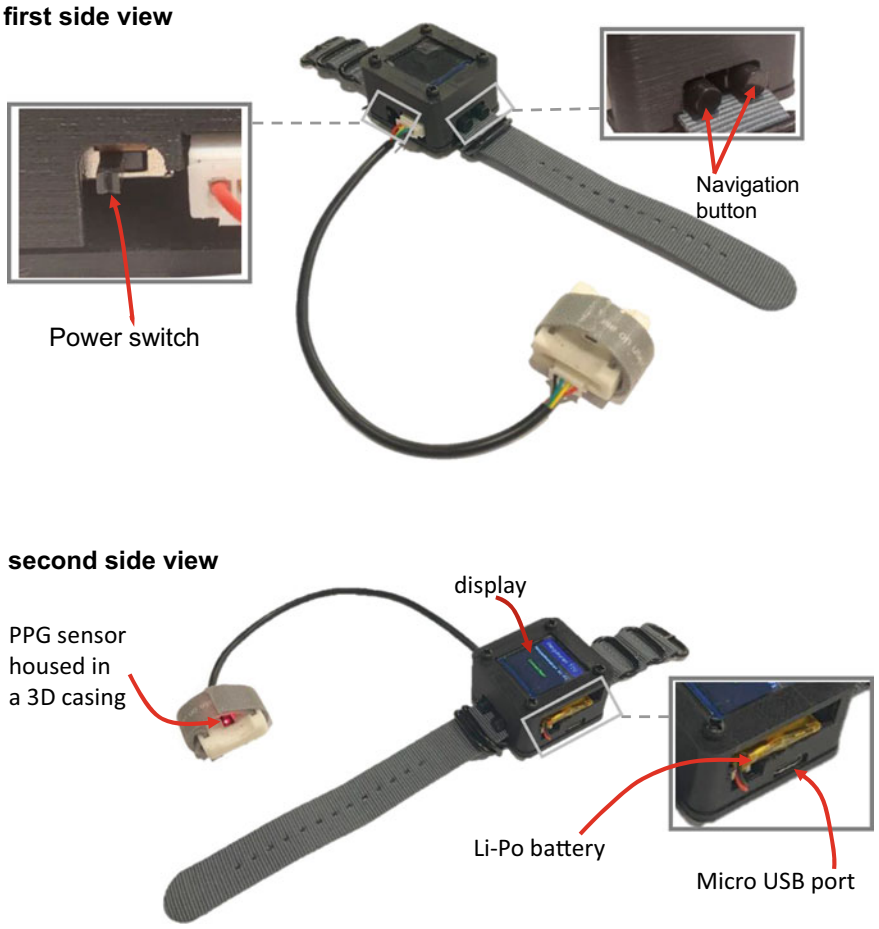


Fig. 4 Wearable respiratory rate measurement device designed with a wristwatch model and wired with a PPG sensor placed on the fingertip

The wearable design of this respiratory rate measuring device is designed with a weight of about 92 g and final dimensions of 4.7 cm × 4.4 cm × 2.7 cm with filament material. On the screen with a size of 4.4 × 4.7 cm, the user can see that this tool is taking measurements for 1 min, as shown in Fig. 7a. On the screen, there is a 1-min countdown display. The results of Heart Rate measurements taken will be displayed on the screen in units of beats per minute (BPM) and Respiration Rate in units of BrPM, as shown in Fig. 7b.

The evaluation performance of the developed algorithm is carried out by calculating the mean difference (MD) value between the actual Respiratory Rate and the predicted Respiratory Rate, as well as the accuracy of testing five subjects with a total of 50 samples from 5 subjects. The (4) and (6) equations define the MD formula

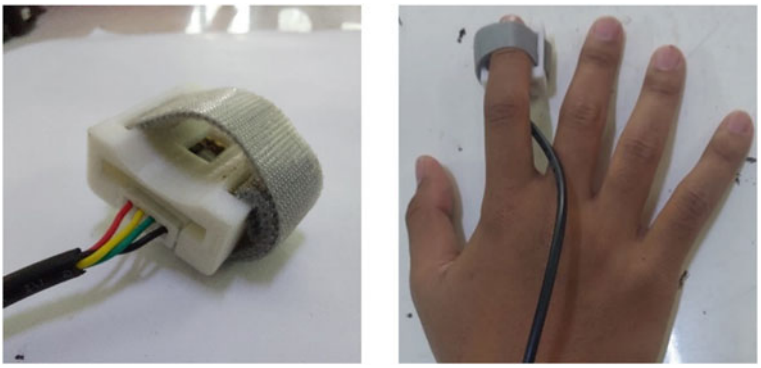


Fig. 5 **a** Application of 3D design of sensor case and Velcro tape as adjustable finger adhesive based on finger size; **b** sensor wearing position on the index finger

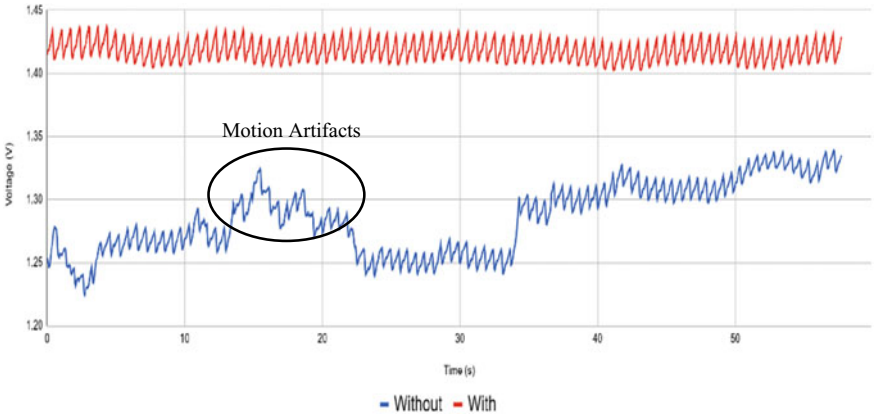


Fig. 6 Comparison of the PPG signal generated by the sensor when not using and using Velcro. The blue signal indicates the presence of noise or motion artifacts

and accuracy.

$$MD = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \tag{4}$$

$$Error = \frac{|y_i - x_i|}{x_i} \times 100 \tag{5}$$

$$Accuracy = 100\% - (error) \tag{6}$$

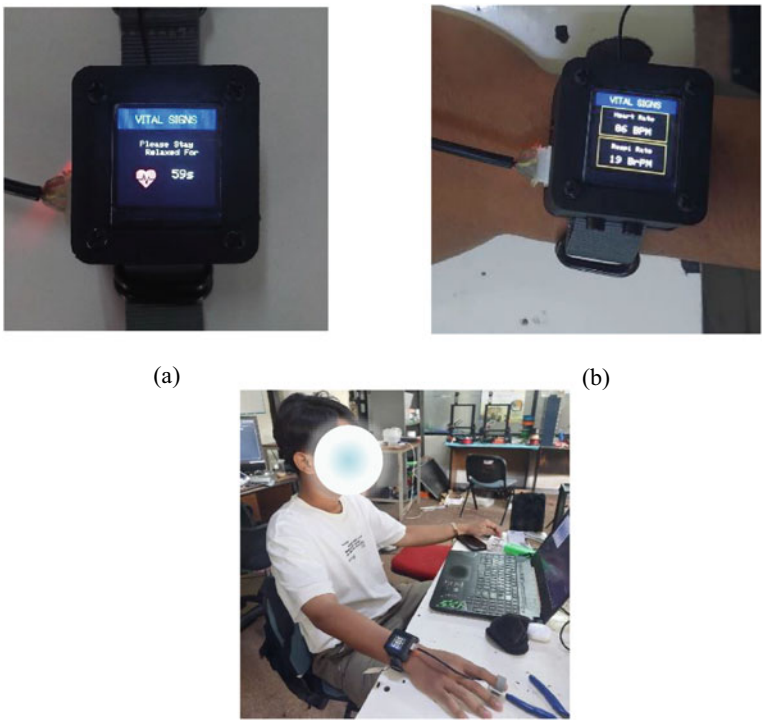


Fig. 7 **a** Display of the device during the measurement. **b** Display of HR and RR measurement results on the device. **c** The subject's sitting position requires you to be calm and relaxed for 1 min of measurement with hands on the table

where $x_1, x_2, \ldots, x_n$ represent the actual respiratory rate and $y_1, y_2, \ldots, y_n$ represent the predicted respiratory rate values obtained from the developed algorithm.

The measurement data from the first subject are shown in Table 1, while the comparison of actual and predicted RR data is in Fig. 8. During the testing phase, the first subject, aged 15 years, had an average actual respiratory rate (RR) of 17.1 BrPM, with an average predicted RR of 17.4 BrPM. A discrepancy between the actual measurements and the predicted respiratory rate was observed. Figure 8 visually demonstrates a remarkable consistency in the comparison graph, except for the 6th measurement point, which deviates slightly. Upon closer inspection of the comparison, the divergence between actual and predicted RR values encompasses a range from 0 BrPM, indicating minimal discrepancy, to 2 BrPM, reflecting the highest variation observed. The maximum accuracy achieved was 100%, indicating perfect agreement, while the minimum accuracy was 88.89%, suggesting a slight deviation between the measurements.

Table 1 The result of RR measurement in subject 1

Data	Actual RR (BrPM)	Predicted RR (BrPM)	Difference (BrPM)	Error rate (%)	Accuracy (%)
1	17	17	0	0.00	100.00
2	17	18	1	5.88	94.12
3	16	17	1	6.25	93.75
4	14	15	1	7.14	92.86
5	20	19	1	5.00	95.00
6	18	20	2	11.11	88.89
7	18	18	0	0.00	100.00
8	17	16	1	5.88	94.12
9	17	17	0	0.00	100.00
10	17	17	0	0.00	100.00

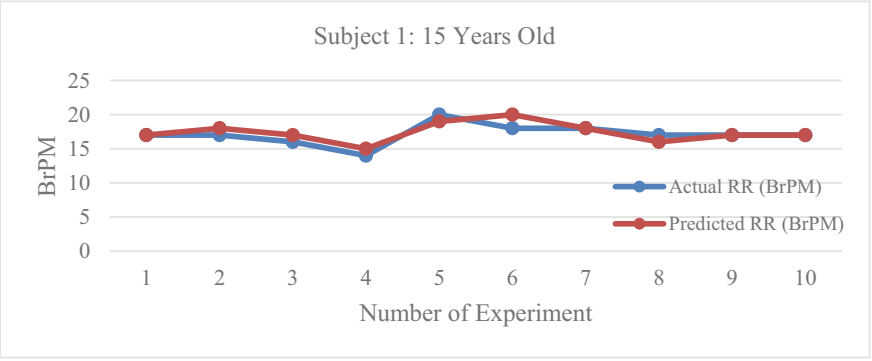


Fig. 8 Comparison of actual and predicted respiratory rate measurement values in subject 1

Moving on to the second subject, aged 48, the average actual RR was 16.3 BrPM, whereas the average predicted RR was 17.3 BrPM, as presented in Table 2. Like the first subject, there was a difference between the actual measurements and the predicted respiratory rate. The comparison between the actual and predicted RR data can be observed in Fig. 9. In this image, a novel pattern appears specifically, and an intriguing consistency emerges, with the actual and predicted RR values consistently lagging by 1 BrPM in 5 out of 10 studies. The differences between the actual and predicted values varied, with a minimum value of 0 BrPM and a maximum value of 3 BrPM in the last experiment. When evaluating the performance, the accuracy assessment provides revealing insights. The accuracy ranges from a substantial 100% precision, indicating precise alignment between model predictions and actual measurements, to a more moderate 81.25% accuracy score in scenarios where deviations are slightly more noticeable.

Table 2 The result of RR measurement in subject 2

Data	Actual RR (BrPM)	Predicted RR (BrPM)	Difference (BrPM)	Error rate (%)	Accuracy (%)
1	16	17	1	6.25	93.75
2	16	17	1	6.25	93.75
3	16	17	1	6.25	93.75
4	15	17	2	13.33	86.67
5	16	17	1	6.25	93.75
6	17	17	0	0.00	100.00
7	17	17	0	0.00	100.00
8	18	18	0	0.00	100.00
9	16	17	1	6.25	93.75
10	16	19	3	18.75	81.25

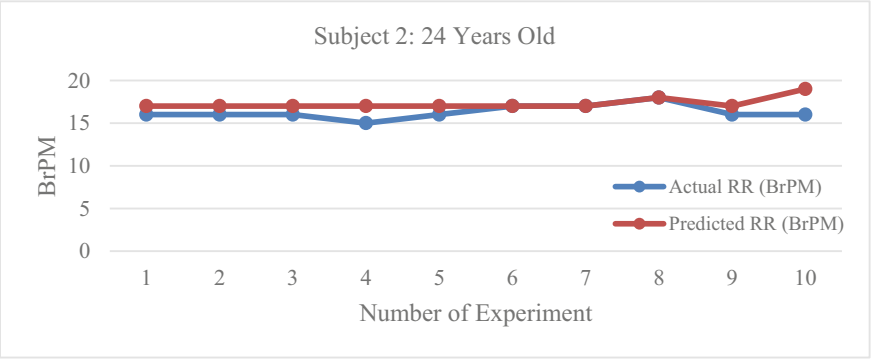


Fig. 9 Comparison of actual and predicted respiratory rate measurement values in subject 2

Table 3 shows data from the measurement findings of subject 3, while Fig. 10 visually synthesizes the actual and predicted respiratory rate (RR) values. In the case of the third subject, aged 29, the mean actual RR was 19.1 BrPM, slightly exceeding the average predicted RR of 18.8 BrPM. The accuracy spectrum ranges from a maximum of 100% precision to a minimum of 85%. Notably, deviations between actual and predicted RR values were observed. As subject three measurement data are compared, there is a fluctuation in the difference between the actual and expected values, with the actual value occasionally being smaller or the predicted value being higher. The range of BrPM between the actual and expected values was 0 to 3 BrPM, highlighting the most significant disparities in RR.

Moving on to the fourth subject, an individual aged 33, and focusing on the output data presented in Table 4. The computed mean value for the actual RR stands at 17.6 BrPM, whereas the average predicted RR, calculated at 18.3 BrPM, gives a discernible variance. A difference was observed between the actual measurements

Table 3 The result of RR measurement in subject 3

Data	Actual RR (BrPM)	Predicted RR (BrPM)	Difference (BrPM)	Error rate (%)	Accuracy (%)
1	20	20	0	0.00	100.00
2	18	20	2	11.11	88.89
3	19	19	0	0.00	100.00
4	19	20	1	5.26	94.74
5	19	18	1	5.26	94.74
6	19	19	0	0.00	100.00
7	20	17	3	15.00	85.00
8	19	19	0	0.00	100.00
9	19	17	2	10.53	89.47
10	19	19	0	0.00	100.00

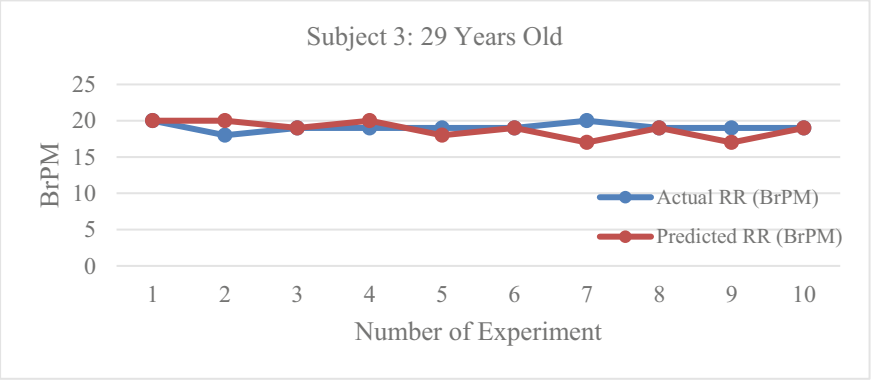


Fig. 10 Comparison of actual and predicted respiratory rate measurement values in subject 3

and the predicted respiratory rate, as shown in Fig. 11. Within this graphical representation, an initial disjunction between the actual RR and predicted RR values is apparent at the onset of the experimental timeline. However, as the experiment progresses, these values converge, reducing the gap between the two. This convergence indicates a gradual alignment between the predictive model's estimations and the actual physiological responses exhibited by the subject. The disparities between the actual values and their corresponding predicted counterparts range from 0 to 3 BrPM. The maximum accuracy was 100%, while the minimum accuracy was 82.35%.

Lastly, the fifth subject, aged 48 and focusing on the output data as presented in Table 5 had an average actual RR of 21 BrPM, while the average predicted RR was 20.3 BrPM. Like the other subjects, a difference existed between the actual measurements and the predicted respiratory rate, as shown in Fig. 12. The actual and predicted values differed from 0 to 3 BrPM. The maximum accuracy achieved

Table 4 The result of RR measurement in subject 4

Data	Actual RR (BrPM)	Predicted RR (BrPM)	Difference (BrPM)	Error rate (%)	Accuracy (%)
1	17	17	0	0.00	100.00
2	17	20	3	17.65	82.35
3	17	19	2	11.76	88.24
4	17	19	2	11.76	88.24
5	19	18	1	5.26	94.74
6	17	17	0	0.00	100.00
7	18	18	0	0.00	100.00
8	18	18	0	0.00	100.00
9	18	18	0	0.00	100.00
10	18	19	1	5.56	94.44

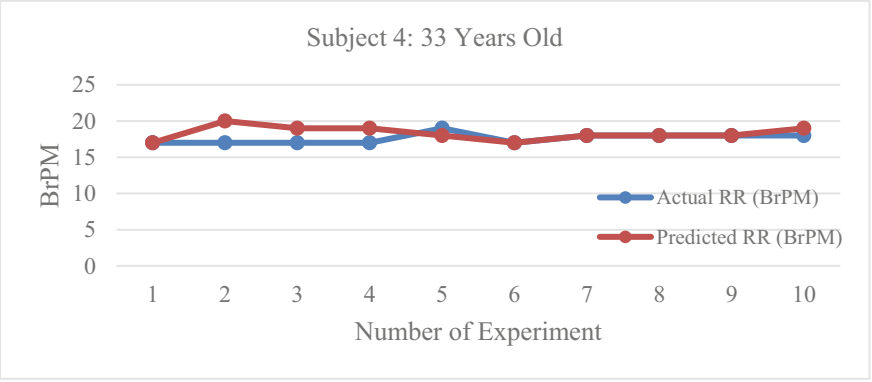


Fig. 11 Comparison of actual and predicted respiratory rate measurement values in subject 4

was 100%, while the minimum accuracy was 85%. Notably, motion artifacts were observed in this subject's measurements, resulting in more differences between the actual and predicted values. This occurred due to the matter speaking during the measurement, introducing noise into the signal [35]. Under normal circumstances, the subject's respiratory rate might have remained relatively stable within a certain range. However, the RR showed fluctuations and changes when the subject suddenly began speaking.

Throughout the 50 tests conducted on the five subjects, the most significant difference observed was 3 BrPM, occurring in four out of five subjects, while the slightest difference was 0. However, these differences are relatively small and acceptable for use in healthy individuals under relaxed conditions.

Table 6 summarizes the MD and accuracy of the respiratory rate predictions tested on five subjects, including 50 samples, regarding manual respiratory rate calculations.

Table 5 The results of RR measurement in subject 5 with the condition of speaking suddenly

Data	Actual RR (BrPM)	Predicted RR (BrPM)	Difference (BrPM)	Error rate (%)	Accuracy (%)
1	25	22	3	12.00	88.00
2	21	21	0	0.00	100.00
3	21	21	0	0.00	100.00
4	22	19	3	13.64	86.36
5	21	20	1	4.76	95.24
6	20	17	3	15.00	85.00
7	20	23	3	15.00	85.00
8	20	18	2	10.00	90.00
9	20	20	0	0.00	100.00
10	20	22	2	10.00	90.00

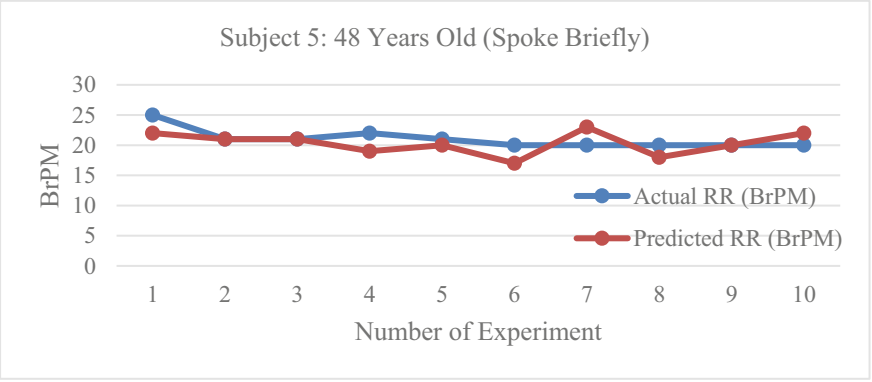


Fig. 12 Comparison of actual and predicted respiratory rate measurement values in subject 5

From testing five subjects with different age ranges and physical conditions, the average accuracy value was 94.32%. Subjects 4 and 5, aged 33 and 26, respectively, achieved the highest level of accuracy among the five individuals tested.

4 Discussion

Based on the summarized findings of the research presented in Table 7, it is evident that the predictive outcomes for respiratory rate values achieved an average of 0.7 ± 0.67 with an average accuracy of 94.32%. The most significant variance in RR values across the dataset was 3 BrPM, while the most minor deviation observed was 0 BrPM. The calculated mean absolute error (MAE) yielded an average value of 1.04

Table 6 Summary of measurement results for five subjects

	Actual RR	Predicted RR	MAE ± STD	Accuracy (%)
Subject 1	17.1	17.4	0.7 ± 0.67	95.9
Subject 2	16.3	17.3	1 ± 0.67	93.7
Subject 3	19.1	18.8	0.9 ± 1.12	95.3
Subject 4	17.6	18.3	0.9 ± 1.17	94.8
Subject 5	21	20.3	1.7 ± 1.41	92.0
Average				94.32

BrPM. By assessing absolute error values, it can be ascertained that the measurement accuracy for respiratory rate quantification exhibited an error margin of 5.68%. This range of accuracy metrics emphasizes the proficiency of the predictive model across different instances, underlining its capacity to generate highly accurate predictions in some cases while maintaining a commendable level of precision in scenarios where deviations are more pronounced.

Factors contributing to these differences may include sudden movements by the subjects, speaking during measurements, or potential biases in the researcher's manual counting of breaths [35–37]. In conclusion, this experiment demonstrates that the tested device can validly measure respiratory rate in healthy individuals under stationary and relaxed conditions. The device's wearable design offers an advantage in convenience and ease of use.

After conducting several tests and data collection, the device's performance was compared to several previous research tools. The comparison was based on the sensor type, application location, processing method, performance, and related display information. The comparison between the proposed study and other studies is presented in Table 7.

Manoni and his team developed a respiratory measurement device using the same type of sensor. However, the device was designed to be placed on the nasal septum, achieving an accuracy of 93% [38]. The developed device did not have a display like the proposed study. The research conducted by Mahardika and his team used a

Table 7 Comparison with past studies

	Sensor	Application area	Processing method	Performance (%)	Displays value
Proposed study	MAX30102	Fingertip	Peak Calculation	94.32	Yes, OLED TFT 1.3-inch
Previous study 1 [38]	MAX30102 LSM6DSM (IMU)	Nasal Septum	PSD PWA	93	No
Previous study 2 [39]	Flex sensor	Abdomen	Curvature abdomen calculation	96.80	Yes, LCD 16 × 2

flex sensor with a respiratory rate calculation method based on the curvature of the sensor attached to the abdomen, achieving a system performance of 96.80% [39]. This device had a display, but it had larger dimensions and was not wearable.

This comparison proves that the peak calculation method using signal processing with a wearable device design improves respiratory rate measurement. The limitation of this device is that users must be in a relaxed seated position without much movement, while the hope is that it can be used in any condition. The implication of this study is the ease of measuring respiratory rate non-invasively with high accuracy in the conditions indicated. This research can increase public awareness of their respiratory rate to anticipate abnormalities immediately.

5 Conclusion

This study aims to develop a non-invasive, wearable, and user-friendly respiratory rate measurement device utilizing PPG signals. The findings show that the peak calculation algorithm can effectively measure the respiratory rate. Measurements performed on five subjects, each undergoing ten measurements, showed an average accuracy rate of 94.32%. The progress in wearable respiratory rate measurement devices signifies a notable stride towards more efficient and accurate healthcare technology, ultimately benefiting medical practitioners and patients. However, further research and validation may be necessary to ensure widespread adoption and integration into standard medical practices. The device can be further developed in future research to measure other vital signs using PPG sensors placed on the fingertips. In addition, the vital sign data obtained can effectively be used for urgency-level classification in a single measurement device.

References

1. Loughlin PC, Sebat F, Kellett JG (2018) Respiratory rate: the forgotten vital sign—make it count! Joint Comm J Qual Patient Safety
2. Guruh IB, Maghfiroh AM, Solichin AA, Bintoro FY (2022) Vital sign monitor based on telemedicine using android application on mobile phone. In: Triwiyanto T, Rizal A, CW (ed) Proceedings of the 2nd international conference on electronics, biomedical engineering, and health informatics. Springer Nature Singapore, Singapore, pp 73–83
3. Nicolò A, Massaroni C, Schena E, Sacchetti M (2020) The importance of respiratory rate monitoring: from healthcare to sport and exercise. Sensors (Switzerland) 20:1–45
4. Institute of Electrical and Electronics Engineers Malaysia Section, Annual IEEE computer conference, IEEE symposium on computer applications and industrial electronics 2015.04.12–14 Langakwi, ISCAIE 2015.04.12–14 Langakwi Respiratory Rate of Photoplethysmogram Signal from Anaesthetic Patients
5. Ren L, Wang H, Naishadham K, Liu Q, Fathy AE (2015) Non-invasive detection of cardiac and respiratory rates from stepped frequency continuous wave radar measurements using the state space method. In: 2015 IEEE MTT-S international microwave symposium, pp 1–4

6. Reuter S, Moser C, Baack M, Falls S (2014) Respiratory distress in the newborn. American Academy of Pediatrics
7. Prinable J, Jones P, Boland D, McEwan A, Thamrin C (2020) Derivation of respiratory metrics in health and asthma. Sensors (Switzerland) 20:1–11
8. Sepriani R (2017) Pengaruh pemberian minuman beroksigen terhadap kemampuan volume Oksigen Maksimal (VO2 Maks). Menssana 2
9. Kumar P, Sharma A, Zimik M, Parashar S, Singh R, Bora J, Patra AK, Mishra M (2020) Measurement of heartbeats for well-being assessment using photoplethysmographic signals. In: Mallick PK, Meher P, Majumder A, Das SK (ed) Electronic systems and intelligent computing. Springer Singapore, Singapore, pp 485–494
10. Bavkar Vandana C, Shinde A (2023) The use of photoplethysmography for blood glucose estimation by noninvasive method. In: Shreesha C, Bandyopadhyay R (ed) Smart sensors measurement and instrumentation. Springer Nature Singapore, Singapore, pp 323–335
11. Ungureanu G-M, Aarts RM, Arends J (2014) Real-time extraction of the respiratory rate from photoplethysmographic signal using wearable devices
12. Hoilett O, Twibell A, Srivastav R, Callihan Linnes J (2018) Kick LL: a smartwatch for monitoring respiration and heart rate using photoplethysmography
13. Shuzan MNI, Chowdhury MH, Chowdhury MEH, Murugappan M, Hoque Bhuiyan E, Arslane Ayari M, Khandakar A (2023) Machine learning-based respiration rate and blood oxygen saturation estimation using photoplethysmogram signals. Bioengineering 10. https://doi.org/10.3390/bioengineering10020167
14. Dong X, Wang Z, Cao L, Chen Z, Liang Y (2023) Whale optimization algorithm with a hybrid relation vector machine: a highly robust respiratory rate prediction model using photoplethysmography signals. Diagnostics 13. https://doi.org/10.3390/diagnostics13050913
15. Boccignone G, D'Amelio A, Ghezzi O, Grossi G, Lanzarotti R (2023) An evaluation of non-contact photoplethysmography-based methods for remote respiratory rate estimation. Sensors 23. https://doi.org/10.3390/s23073387
16. Javaid M, Haleem A, Pratap Singh R, Suman R, Rab S (2022) Significance of machine learning in healthcare: features, pillars and applications. Int J Intell Netw 3:58–73. https://doi.org/10.1016/j.ijin.2022.05.002
17. Goecks J, Jalili V, Heiser LM, Gray JW (2020) How machine learning will transform biomedicine. Cell 181:92–101
18. Weissler EH, Naumann T, Andersson T, Ranganath R, Elemento O, Luo Y, Freitag DF, Benoit J, Hughes MC, Khan F, Slater P, Shameer K, Roe M, Hutchison E, Kollins SH, Broedl U, Meng Z, Wong JL, Curtis L, Huang E, Ghassemi M (2021) The role of machine learning in clinical research: transforming the future of evidence generation. Trials 22
19. Přibil J, Přibilová A, Frollo I (2020) Comparative measurement of the PPG signal on different human body positions by sensors working in reflection and transmission modes
20. Sulistyowati A (2018) Pemeriksaan Tanda Tanda Vital. Akademi Keperawatan Kerta Cendekia Sidoarjo
21. Kang M, Shin S, Kang KH, Kim YT (2020) Multi bio-signal based algorithm using EMD and FFT for stress analysis. In: 2020 IEEE international conference on consumer electronics (ICCE), pp 1–4
22. Park C, Shin H, Lee B (2017) Blockwise PPG enhancement based on time-variant zero-phase harmonic notch filtering. Sensors (Switzerland) 17. https://doi.org/10.3390/s17040860
23. Mohd Sani NH, Mansor W, Lee KY, Ahmad Zainudin N, Mahrim SA (2015) Determination of heart rate from photoplethysmogram using Fast Fourier transform. In: 2015 international conference on BioSignal analysis, processing and systems (ICBAPS), pp 168–170
24. Park J, Seok HS, Kim SS, Shin H (2022) Photoplethysmogram analysis and applications: an integrative review. Front Physiol 12
25. An X, Stylios GK (2020) Comparison of motion artefact reduction methods and the implementation of adaptive motion artefact reduction in wearable electrocardiogram monitoring
26. Lee J-W, Yoon S-J, Kim K-H (2021) Zero phase filter design and system study for improved handling of ankle PPG signals. Trans Korean Inst Electr Eng 70:243–248. https://doi.org/10.5370/KIEE.2021.70.1.243

27. Drummond GB, Fischer D, Arvind DK (2020) Current clinical methods of measurement of respiratory rate give imprecise values. ERJ Open Res 6. https://doi.org/10.1183/23120541.00023-2020
28. Koziel A, Jarmuszkiewicz W (2017) Hypoxia and aerobic metabolism adaptations of human endothelial cells. Pflugers Arch 469:815–827. https://doi.org/10.1007/s00424-017-1935-9
29. Moraes Filho AA, Schreiber G, Sieg JA, Much MD, Moura Bartoski V, Marcon C (2023) Methods to estimate respiratory rate using the photoplethysmography signal. In: 16th international joint conference on biomedical engineering systems and technologies (BIOSTEC 2023) - Volume 5: Healthinf, pp 445–452
30. Mohd Sani NH, Mansor W, Lee KY, Ahmad Zainudin N, Mahrim SA (2015) Determination of heart rate from photoplethysmogram using Fast Fourier transform. In: 2015 international conference on BioSignal analysis, processing and systems, ICBAPS 2015. Institute of Electrical and Electronics Engineers Inc., pp 168–170
31. He X, Goubran RA, Liu XP (2014) Secondary peak detection of PPG signal for continuous cuffless arterial blood pressure measurement. IEEE Trans Instrum Meas 63:1431–1439. https://doi.org/10.1109/TIM.2014.2299524
32. Yulian R, Suprianto B (2017) Rancang Bangun Photoplethysmography (PPG) Tipe Gelang Tangan untuk Menghitung Detak Jantung Berbasis Arduino
33. Ishikawa T, Hyodo Y, Miyashita K, Yoshifuji K, Komoriya Y, Imai Y (2017) Wearable motion tolerant PPG sensor for instant heart rate in daily activity. In: BIOSIGNALS 2017 - 10th international conference on bio-inspired systems and signal processing, proceedings; Part of 10th international joint conference on biomedical engineering systems and technologies, BIOSTEC 2017. SciTePress, pp 126–133
34. Finnegan E, Davidson S, Harford M, Watkinson P, Tarassenko L, Villarroel M (2023) Features from the photoplethysmogram and the electrocardiogram for estimating changes in blood pressure. Sci Rep 13. https://doi.org/10.1038/s41598-022-27170-2
35. Almarshad MA, Islam MS, Al-Ahmadi S, Bahammam AS (2022) Diagnostic features and potential applications of PPG signal in healthcare: a systematic review. Healthcare (Switzerland) 10
36. Ghamari M (2018) A review on wearable photoplethysmography sensors and their potential future applications in health care. Int J Biosens Bioelectron 4. https://doi.org/10.15406/ijbsbe.2018.04.00125
37. Pi I, Pi I, Wu W (2021) External factors that affect the photoplethysmography waveforms. SN Appl Sci 4:21. https://doi.org/10.1007/s42452-021-04906-9
38. Manoni A, Loreti F, Radicioni V, Pellegrino D, Della TL, Gumiero A, Halicki D, Palange P, Irrera F (2020) A new wearable system for home sleep apnea testing, screening, and classification. Sensors (Switzerland) 20:1–26. https://doi.org/10.3390/s20247014
39. Mahardika IKEG, Hamzah T, Rahmawati T, Soetjiatie L (2019) Measuring respiration rate based android. Indonesian J Electron Electromed Eng Med Inf 1:39–44. https://doi.org/10.35882/ijeeemi.v1i1.7

Automatic Obstructive Sleep Apnea Identification Using First Order Statistics Features of Electrocardiogram and Machine Learning

Aida Noor Indrawati, Nuryani Nuryani, Wiharto Wiharto, Diah Kurnia Mirawati, and Trio Pambudi Utomo

Abstract Obstructive Sleep Apnea (OSA), or cessation of breathing during sleep, is caused by a blocked upper airway. OSA is one of the causes of sudden death caused by heart failure during sleep. Given the impact of OSA, early detection is therefore necessary. This article presents an innovation in OSA detection using first-order statistical features of an electrocardiogram (ECG) and machine learning. The first-order statistical features were obtained from the RR interval of the ECG single-lead. Various statistical features were evaluated for OSA detection. Moreover, machine learning techniques such as Linear Discriminant Analysis (LDA), Artificial Neural Networks (ANN), K-Nearest Neighbors (K-NN), and Support Vector Machines (SVM) were investigated for OSA detection. The accuracy, sensitivity, and accuracy values of each classification are the performance outcomes of machine learning. The proposed OSA detection method was validated using clinical data from patients diagnosed with sleep apnea. When using KNN for OSA detection, the results showed an accuracy of 83.91%, sensitivity of 92.31%, and specificity of 67.19%. Using SVM, the OSA detection demonstrated an accuracy of 83.75%, sensitivity of 92.97%, and specificity of 65.80%. The ANN method for OSA detection yielded an accuracy of 80.83%, sensitivity of 92.67%, and specificity of 59.25%. Lastly, OSA detection using LDA exhibited an accuracy of 79.27%, sensitivity of 82.27%, and specificity of 73.29%. From the results, it was concluded that the KNN classification, with all its features, provided the best performance for OSA detection, achieving the highest accuracy of 83.91%.

A. N. Indrawati · N. Nuryani (✉) · T. P. Utomo
Department of Physics, Sebelas Maret University, Surakarta, Indonesia
e-mail: nuryani@mipa.uns.ac.id

W. Wiharto
Department of Informatic, Sebelas Maret University, Surakarta, Indonesia

D. K. Mirawati
Department of Medical Sciences, Faculty of Medicine, Sebelas Maret University, Surakarta, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_11

Keywords Electrocardiography · Sleep Apnea · Obstructive · Machine learning

1 Introduction

Obstructive Sleep Apnea (OSA) is a breathing disorder that occurs while you sleep and is characterized by brief breathing pauses. Upper airway blockage is what makes OSA happen. Despite repeated attempts to breathe, the airway is obstructed at the level of the tongue or soft palate, preventing air from entering the lungs. This typically causes a drop in blood oxygen saturation and wakefulness from sleep to breathe [1]. Low sleep quality in OSA patients can result in poor concentration, memory loss, slow responses, and depression [2]. Additionally, OSA may result in poorer daytime cognitive performance, increasing the risk of motor vehicle accidents. Undiagnosed OSA is now considered an important risk factor for the development of cardiovascular disease (e.g. hypertension, stroke, heart failure, left ventricular hypertrophy, acute coronary syndrome) [3].

Given the devastating effects of OSA, early detection crucial. Polysomnography (PSG) is a method used to diagnose OSA. PSG tests are conducted over the course of an entire night at a sleep center or sleep lab. A sleep specialist personally evaluated the PSG monitoring data in accordance with the recommendations of the American Academy of Sleep Medicine (AASM) [4]. Although the diagnosis of OSA made using PSG data is thorough and precise, there are certain evaluation restrictions with PSG. The manual assessment process makes PSG examinations time-consuming, there are lengthy wait periods, few large hospitals have PSG technology, and only a small number of patients are monitored every night. Additionally, PSG monitoring is costly and uncomfortable [4]. PSG examination is cumbersome since it diagnoses OSA using multiple techniques at once. These instruments come from non-invasive sensors for chest and abdominal breathing, oxyhemoglobin levels, tracheal sounds, electroencephalography (EEG), electrooculography (EOG), electromyography (EMG), and electrocardiography (ECG) [5].

Due to the many shortcomings of testing with PSG, many OSA detection methods have been developed. One of the OSA detection methods currently being developed is using only a single lead electrocardiogram (EKG). EKG is a tool to record the electrical activity of the heart. The working principle of the EKG is to measure the biopotential differences measured on the outside of the body [6]. Some machine learning has been carried out for OSA detection, including that done by Indrawati et al. (2022) regarding OSA classification using various Machine learning methods [7]. Rajesh et al. (2021) regarding OSA detection with HRV signals and decomposition with discrete wavelet transform (DWT) [8]. Zhu et al. (2021) regarding OSA classification using the Support Vector Machine (SVM) method, K-Nearest Neighbor (K-NN), logistic regression, and Random Forest (RF) [2]. Gütüler et al. (2014) have conducted research on the diagnosis of obstructive sleep apnea using Neural Networks (NN) classification [9]. Magnusdottir and Hilmisson (2017) have conducted research on OSA detection using the Cardio Pulmonary Coupling (CPC)

and Cyclic Variation of Heart Rate (CVHR) features [10]. Hassan and Haque (2017) have conducted research on the detection of Obstructive Sleep Apnea (OSA) with statistical and spectral features using Bootstrap [11]. Sharma and Sharma (2020) have conducted sleep apnea research using various decomposition modes [12].

Until now the OSA detection method has been used to get the best performance. This is what underlies this research on OSA detection with RR interval features using first-order feature variations. In this study, various machine learning methods were used, namely Linear Discriminant Analysis (LDA), Artificial Neural Networks (ANN), K-Nearest Neighbors (K-NN), and Support Vector Machines (SVM).

This article heralds a multi-dimensional contribution to the field of Obstructive Sleep Apnea (OSA) detection. At its core, the introduction of "First Order Statistics Features" stands as a groundbreaking approach, potentially enhancing our ability to understand and detect OSA. These features, rooted in the primary statistical properties of datasets, offer an innovative lens through which OSA's intricate patterns might be discerned, thereby facilitating improved early detection rates. This article marries these First Order Statistics Features with four esteemed machine learning algorithms: SVM, LDA, ANN, and k-NN. This integration not only underscores the versatility and adaptability of these features but also establishes a foundational benchmark for future research endeavors. The narrative takes an even deeper dive into the realm of tailored machine learning by emphasizing the optimization of the SVM model. By specifically fine-tuning SVM for OSA detection using these new features as input, the article pushes the boundaries of model efficiency and accuracy. Such an optimization process acts as a beacon, illuminating the importance of customizing machine learning models to address unique challenges head-on. All in all, the article's layered contributions signify a monumental stride in OSA detection methodologies, bridging gaps in knowledge and lighting the path for subsequent innovations.

2 Materials and Method

The Apnea ECG data used in this investigation can be obtained from the physionet.org website. Severe OSA (class A), moderate OSA (class B), and normal OSA (class C) make up the three categories of the data set that was employed [13]. The dataset utilized to test the algorithm for diagnosing OSA consisted of 35 patients, 20 patients in patient group A, 5 patients in patient group B, and 10 patients overall. Class A consisted of AHI $\geq$ 10 per hour (More than 100 min of apnea on all recordings). Class B consisted of 10 > AHI $\geq$ 5 per hour (there were 5–99 min of apnea on all recordings). Class C consisted of AHI < 5 per hour (less than 5 min of apnea on all recordings) [13]. This research was conducted by following the appropriate stages in Fig. 1.

Data from each patient's peak R ECG was used in this investigation. Then, all patients underwent segmentation for 60 s. Following segmentation, the RR interval is looked for. With $i = 1, 2, 3, \ldots, n$, the RR interval is calculated by subtracting wave

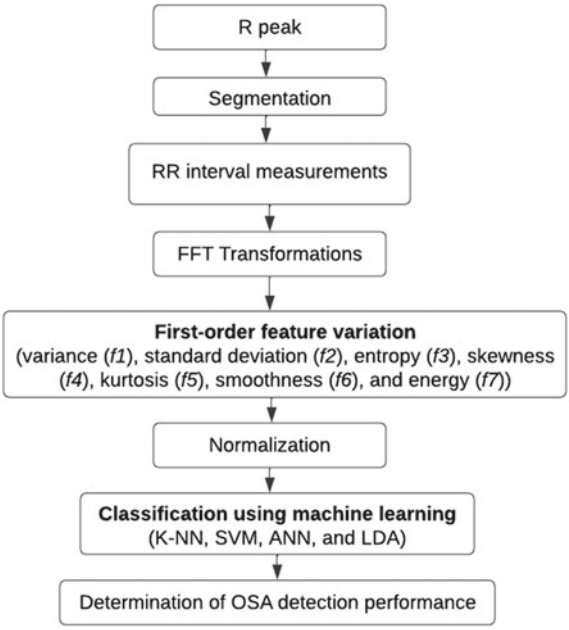


Fig. 1 OSA detection system in general

$R(i+1)$ from wave $R(i)$. The next stage is the Fast Fourier Transform (FFT) process for each segment. FFT is an algorithm used to represent signals as frequencies in the time domain [14]. The FFT of a signal $x(n)$, can be calculated by [14]:

$$\begin{aligned} X(k) &= \sum_{n=0}^{N-1} x(n) W_N^{nk} \\ &= \sum_{n\,genap} x(n) W_N^{nk} + \sum_{n\,ganjil} x(n) W_N^{nk} \\ &= \sum_{m=0}^{(\frac{N}{2})-1} x(2m) W_N^{2mk} + \sum_{m=0}^{(\frac{N}{2})-1} x(2m) W_N^{2mk} \end{aligned} \tag{1}$$

where k is an index in the frequency domain $= 0, 1, \ldots, N-1$ and n is an index in the time domain $= 0, 1, \ldots, N-1$.

The calculation of statistical features is done after getting the output from the FFT. The statistical features used in this study are first-order feature variations, namely variance ($f1$), standard deviation ($f2$), entropy ($f3$), skewness ($f4$), kurtosis ($f5$), smoothness ($f6$), and energy ($f7$). as in Table 1 [15].

After calculating the statistics, the next step is normalization. Normalization is carried out because the statistical data results are irregular. Normalization is carried out so that the data is at a predetermined scale or range [16]. In this study, Eq. 2 was used to do Min–Max normalization (Xn).

Table 1 Variation of first-order features in statistical calculations

Equation	Description
$f1 = \sum^{n-1}\left(1 - \frac{\sum X(k)}{n}\right)^2 X[i]$	$\sum X(k)$= the total amount of data n = amount of data $X[i]$ = value i
$f2 = \sqrt{\frac{\sum^{n-1}(X(k)_i - X(k)_{ave})^2}{n-1}}$	
$f3 = -\sum^{n-1} X[i] \log 2X[i]$	
$f4 = \sum^{n-1}\left(1 - \frac{\sum X(k)}{n}\right)^3 X[i]$	
$F5 = \sum^{n-1}\left(1 - \frac{\sum X(k)}{n}\right)^4 X[i]$	
$f6 = 1 - \frac{1}{1+f1}$	
$F7 = \sum^{n-1}(X[i])^2$	

$$X_n = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{2}$$

where x is the complex QRS width and gradient, $x_{\min}$ is the minimum data, and $x_{\max}$ is the maximum data [16].

The normalized data is then divided into two, parts: 20% testing data and 80% training data. The classification in this study is by using several machine learning algorithms, namely K-NN, SVM, ANN, and LDA.

2.1 K-Nearest Neighbor (K-NN)

One technique for classifying data is the K-Nearest Neighbor (K-NN) algorithm. The underlying idea behind K-NN is to categorize data based on how similar it is to other data. Finding the shortest path between the evaluation data and the K nearest neighbors in the training data is how K-NN operates [17]. The following equation can be used to determine Euclidean distance, which is used by the K-NN function.

$$D(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2} \tag{3}$$

where p and q are the comparison subjects, and there are n attributes [17].

The supervised learning model incorporates the K-NN classification, which means that the population database used has already been assigned to a specific class. The way this method operates is by looking for as many K-objects in the database as possible that are the same size as the item under test. The features used determine the proximity measurement [17].

2.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a learning technique using pattern recognition. This classification is processed using input data input and output data in the form of expected targets, so it is often referred to as directed learning (supersived learning). The working principle of SVM is Structural Risk Minimization (SRM), whose function is to find the best hyperplane so that the two classes in the input space can be separated [18]. In general, the SVM function can be calculated by Eq. 4.

$$f_{SVM}(\mathbf{x}) = w^T \phi(\mathbf{x}) + b \tag{4}$$

The parameters w and b respectively represent the weight and bias vectors which can be determined in the training process by minimizing the function, and $\phi(.)$ is used to map the input vector x into a higher dimensional space, so that it is easily separated by a linear hyperplane [19].

2.3 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is a computational model of an artificial intelligence that is designed according to the biological nervous system. ANN is an adaptive system, meaning that this system can change its shape according to the internal and external information obtained during the learning process. ANN is used to solve a particular problem such as classification or pattern recognition, using a computer program [20]. Where ANN is an artificial intelligence that represents brain cells, namely neurons. The four elements of a neuron, namely the activation function, receptors, summators, and output elements. Backpropagation ANN was employed in this study. ANN To alter a weight that can be connected to the neurons in the hidden region, backpropagation, a learning technique, is typically employed with many layers [21].

In the back-propagation phase, the actual output of the network is compared with the desired output. Error calculation is done by using Eq. 5 [22].

$$dj = (tj-)'(Sj) \tag{5}$$

where Sj is the weighted sum of the inputs to neuron j, tj is the expected output from neuron j, yj is the actual output from neuron j,$'$ (S) is the derivative of the activation function (S) at point x.

2.4 Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is a discriminant analysis that uses multivariate statistics related to separating or classifying a group of objects into a predefined group. In an effort to recognize an object, this method finds groups whose values are able to separate objects whose characteristics are known. The purpose of classifying objects in this way is so objects can be separated into two or more groups. Discriminant analysis is a statistical technique that can be used to differentiate explanatory variables and response variables. The explanatory variable here is in the form of quantitative data while the response variable here is in the form of qualitative data. The assumptions on discriminant analysis in each group are the independent variables with multivariate normal distribution. Whereas for each group, the independent variable covariate matrix is the same [23]. The function of LDA can be seen in Eq. 6 [24].

$$f_i = \mu_i C^{-1} x_k^T - \frac{1}{2}\mu_i C^{-1} \mu_i^T + \ln(Pi) \tag{6}$$

where f_i is the discriminant function of the ith class, μ_i is the average value of each matrix for each class, C^{-1} is the inverse of the group covariance matrix, x_k^T is the transpose of the test data matrix, μ_i^T is the average transpose of the i-th class, Pi is the probability of the i-th class appearing.

Performance assessment comes after classification using K-NN, SVM, ANN, and LDA. Accuracy, specificity, and sensitivity were three characteristics that this study used to evaluate performance. The capacity to accurately identify OSA and normal is known as accuracy. The capacity to accurately identify the norm is known as specificity. Sensitivity is the capacity for accurate OSA diagnosis [25].

The equations for finding specificity, sensitivity, and accuracy values are shown in Eqs. 7, 8 and 9 respectively [25].

$$\text{Specificity} = \frac{TN}{TN + FP} \times 100\% \tag{7}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100\% \tag{8}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{9}$$

FN is False Negative, TN is True Negative, TP is True Positive, and FP is False Positive.

3 Result

In this study, it was started by collecting Apnea ECG data from the Apnea-ECG Database [18]. The data obtained is then processed using the FFT algorithm. Statistical feature calculations are performed using first-order feature variations using the output from the FFT. Variations of the first order features used are variance ($f1$), standard deviation ($f2$), entropy ($f3$), skewness ($f4$), kurtosis ($f5$), smoothness ($f6$), and energy ($f7$). The next stage is normalization. Classification using machine learning is used after normalization. Some machine learning, namely K-NN, SVM, ANN, and LDA. In this study the performance of the detection system was expressed using specificity, sensitivity and accuracy.

In this study the classification used uses all features for K-NN, SVM, ANN, and LDA inputs. This is to find out and compare the performance of the best machine learning algorithms for OSA detection. It can be seen in Table 2 which is the performance of all features ($f1$, $f2$, $f3$, $f4$, $f5$, $f6$, and $f7$) using various classifications.

In Table 2 it can be seen that the highest accuracy is by using the KNN classification, which is equal to 83.91%. The highest sensitivity using the SVM classification is 92.97%. The highest specificity using the LDA classification is 73.29%. KNN classification produces the best performance outcomes overall since it has the highest accuracy value. This is so that the computer can correctly identify OSA and normal if accuracy is high.

In this research, variations of features and single features will be carried out, this is done to find out the best performance when variations of features or single features are carried out. For single features used with first-order features used, namely variance ($f1$), standard deviation ($f2$), entropy ($f3$), skewness ($f4$), kurtosis ($f5$), smoothness ($f6$), and energy ($f7$) can be use $f1, f2$, until with $f7$. While the variation of the features used are single features and combinations of the features used. In this study, KNN classification will be used to determine the performance of single features and feature variations. Performance results with KNN classification for single features can be seen in Table 3.

In Table 3 it can be seen that the highest accuracy is by using the variance feature ($f1$) which is equal to 77.63%. The highest sensitivity using the energy feature ($f7$) is 93.27%. The highest specificity is by using the entropy feature ($f3$), which is 59.88%. Meanwhile, the lowest accuracy and sensitivity using the skewness feature

Table 2 Shows the performance results for all characteristics using different machine learning techniques

Classification	Specificity (%)	Sensitivity (%)	Accuracy (%)
KNN	67.19	92.32	83.91
SVM	65.80	92.97	83.75
ANN	59.25	92.67	80.83
LDA	73.29	82.27	79.27

Table 3 Single feature performance results with KNN

Feature	Specificity (%)	Sensitivity (%)	Accuracy (%)
$f1$ (variance)	48.14	92.43	77.63
$f2$ (standard deviation)	48.02	92.25	77.47
$f3$ (entropy)	59.88	83.71	75.75
$f4$ (skewness)	45.63	77.34	66.75
$f5$ (kurtosis)	42.04	79.27	66.83
$f6$ (smoothness)	48.14	92.37	77.59
$f7$ (energy)	33.89	93.27	73.43

($f4$) is 66.75% and 77.34% respectively. The lowest specificity is obtained by using the energy feature ($f7$), which is 33.89%.

In research, KNN classification will be used to determine the performance of feature variations. Performance results with KNN classification for feature variations can be seen in Table 4.

In Table 4 it can be seen that the highest accuracy and specificity by using all the features are 83.91% and 67.19%. The highest sensitivity using the feature variance ($f1$) and standard deviation ($f2$) is 92.55%. The lowest accuracy and specificity by using a single feature, namely the variance feature (f1) is 77.63% and 48.14%. It can be seen in Table 4, that the more features used the greater the accuracy, and the greater the specificity.

In this study, K = 15 is used in the KNN algorithm for all the performance obtained. Several variations of K have been carried out and the best performance was obtained, namely when K = 15, it is solid as seen in Fig. 2.

The best accuracy is achieved when K = 15, as shown in Fig. 2. But when K = 1, accuracy is at its lowest. When K = 4, accuracy will significantly rise from K = 1, which is the starting point. The accuracy value is ambiguous for K values ranging from 5 to 30.

This study used several features to perform SVM classification as well. In Table 5, you can observe the performance outcomes of SVM classification for different features.

Table 4 Performance results of feature variations with KNN

Variation features	Specificity (%)	Sensitivity (%)	Accuracy (%)
$f1$	48.14	92.43	77.63
$f1+f2$	49.33	92.55	78.35
$f1+f2+f3$	54.57	91.06	79.07
$f1+f2+f3+f4$	63.46	91.42	82.23
$f1+f2+f3+f4+f5$	62.36	91.60	81.99
$f1+f2+f3+f4+f5+f6$	63.34	91.66	82.35
$f1+f2+f3+f4+f5+f6+f7$	67.19	92.31	83.91

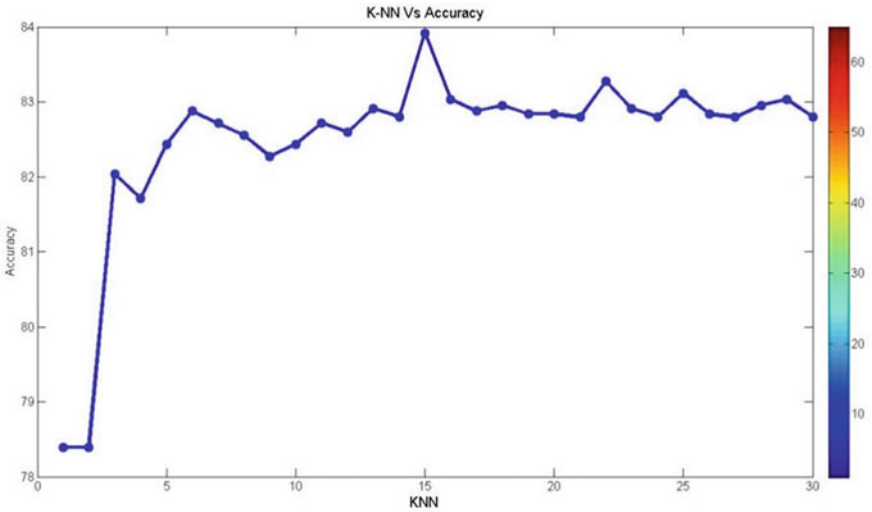


Fig. 2 Graph of detection system accuracy using KNN with variations in k values

Table 5 The performance results of the detection system employing SVM with a range of features

Variation features	Specificity (%)	Sensitivity (%)	Accuracy (%)
$f1$	34.67	98.91	77.11
$f1+f2$	38.09	97.40	77.27
$f1+f2+f3$	42.92	96.91	78.59
$f1+f2+f3+f4$	60.85	93.52	82.43
$f1+f2+f3+f4+f5$	62.26	93.22	82.71
$f1+f2+f3+f4+f5+f6$	62.26	92.91	82.81
$f1+f2+f3+f4+f5+f6+f7$	65.80	92.97	83.75

In Table 5 it can be seen that the highest accuracy and specificity by using all the features are 83.75% and 65.80%. The highest sensitivity is obtained by using a single feature, namely the variance ($f1$), which is equal to 98.91%. The lowest accuracy and specificity by using a single feature, namely the Variance feature ($f1$) is 77.11% and 34.67%. The more features used, the greater the accuracy, the greater the specificity.

In this study, SVM RBF Carnel was used with C = 60, gamma = 100 for all the per performances obtained. Several variations of gamma and C have been carried out, and the best performance was obtained when C = 60, gamma = 100, which is dense, as shown in Fig. 3.

In Fig. 3, the white box with the highest accuracy, which is 83.79% and has a gamma range of 100–250 and a C range of roughly 10–100, can be seen. With a gamma range > 500 and a range C > 500, the lowest accuracy displayed in the

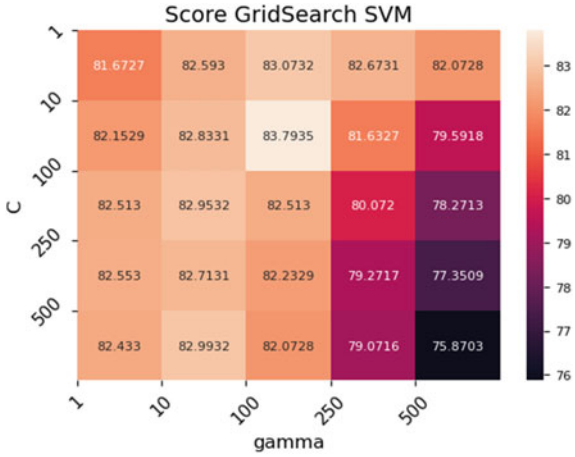


Fig. 3 Accuracy of SVM mapping accuracy with gamma and C variations

black box has an accuracy of 75.87%. A program's accuracy can tell you whether it's effective or not. Because accuracy displays OSA and normal conditions clearly. Low accuracy is caused by a variety of variables, one of which being potential noise in the ECG recording. This study's findings suggest that more features improve performance, whereas fewer characteristics produce smaller performance results.

4 Discussion

In Table 6 you can see various features, classifications and performance results obtained from each study. This study has an accuracy value that is not very good compared to other studies but has quite good sensitivity. This is due to the selection of different features and different classifications.

In this study, the data used were peak R data from each patient. Segmentation was performed every 60 s for each patient. This was also done by Mustofa et al. regarding the automatic identification of OSA using the Convolutional Neural Network (CNN) [26].

This study focuses on the OSA detection technique, which continues to be of significant concern to the medical community. In this study, several approaches have been tested in an effort to find the best one for OSA identification. The first-order characteristics variance ($f1$), standard deviation ($f2$), entropy ($f3$), skewness ($f4$), kurtosis ($f5$), smoothness ($f6$), and energy ($f7$) were used in this work as variations. In order to determine which features work best for OSA detection, both a single feature and a wide range of features are employed.

To ensure that the statistical computation data in this study were on a predetermined scale or range with a range of 0 to 1, Min–Max normalization was used [27].

Table 6 Comparison of the model trained for this study with the work available in the literature for detection OSA

Feature	Classification	Performance	Feature
Mostafa et al. [26]	Segmented SpO2 and HRV combination every 1, 3, and 5 min	CNN	Acu 94%, Sen 92%, spe 96%
Li et al. [4]	Complex QRS detection, using Tomskin pan and band pass filter	FNN, LDA, SVM, Complex tree, RUSBoosted trees, and LR	Acu 97.8%, sen 98.6%, spe 93.9%,
Rajesh et al. [8]	HRV segmented for 1 min, subband decomposition with DWT	LDA, KNN, SVM, and RF	Acu 89.6 ± 0.14 with RF
Zhu et al. [25]	Feature RRI and S122 are segmented by 1 min, bandpass filter is used, various feature statistics are sought	SVM, KNN, LR, and RF	Acu 96.6%, Sen 93.8%, Spe 98.4% with RF
This study	RRI uses FFT	LDA, KNN, NN, and SVM	KNN with acu 83.9%, sen 92.3% and spe 67.2%

After normalization, the data will be split into training and testing data. The performance of a single characteristic and its variants should then be examined. Calculating specificity, sensitivity, and accuracy three parameters leads to the determination of this performance.

In this study, the best performance is by using the KNN algorithm, then SVM. The highest accuracy with the KNN algorithm using all features is 83.91%. Meanwhile, the SVM algorithm has the highest accuracy using all features, namely 83.75%. For classification using SVM it is better than NN obtained by Nguyen et al. for online detection of Sleep Apnea [28]. SVM performance which is better than K-NN is also described by Fatimah and Binish et al. regarding the comparison of several machine learning methods for detecting sleep apnea [28]. SVM (Support Vector Machine) performs well because it is based on the principle of a hyperplane that separates training examples with maximum margin [29]. The use of SVM in classification problems is often considered as pattern recognition using SVM [30]. On the other hand, the performance of Neural Networks (NN) is less satisfactory, which is due to its three components consisting of input layer, hidden layer, and output layer [31]. Meanwhile, the lowest performance is obtained by using the LDA algorithm. This is because the LDA algorithm works by grouping data determined by boundary lines (straight lines) obtained from linear equations [32]. Different from research by Marcos et al. has higher LDA performance than KNN for OSA detection [33].

Even with just one feature, performance is still not at its best. So that the final performance could be even better, many feature combinations were tested in this study. The accuracy and sensitivity will be better if more features are used, according to the performance results. As more features are included, however, the specificity declines.

This research can provide benefits, including contributing to the medical field to detect OSA automatically. Besides that, it provides a wider application of machine learning for OSA detection. Obtain a design that can be applied in the medical world for OSA classification.

The limitation of this research is that the data in this study only comes from the Massachusetts Institute of Technology Beth Israel Hospital (MIT-BIH) which has an ECG Apnea database. The method developed in this research has not provided very good performance, so further research needs to be carried out. It is hoped that future studies will provide better performance than previous studies.

5 Conclusion

In this study, we aimed to develop an efficient OSA detection system utilizing the ECG RR interval. We examined first-order statistical features including variation ($f1$), standard deviation ($f2$), entropy ($f3$), skewness ($f4$), kurtosis ($f5$), smoothness ($f6$), and energy ($f7$). Four distinct classifiers, specifically SVM, ANN, K-NN, and LDA, were employed for OSA detection. The most effective method for OSA detection was classification using KNN with $k = 15$ and incorporating all features ($f1$–$f7$). This approach achieved an accuracy, sensitivity, and specificity of 83.91%, 92.31%, and 67.19%, respectively. We hope that upcoming studies will explore alternative features to discern which are most effective in distinguishing OSA from normal conditions. Additionally, we advocate for future research to utilize primary data directly from patients, which could offer a more straightforward alternative for evaluating OSA patients.

References

1. de Chazal P et al (2003) Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea. IEEE Trans Biomed Eng 50(6):686–696
2. Zhu J et al (2022) Detection of sleep apnea from electrocardiogram and pulse oximetry signals using Random Forest. Appl Sci 12(9):4218–4218
3. Khandoker AH, Karmakar CK, Palaniswami M (2009) Automated recognition of patients with obstructive sleep apnoea using wavelet-based features of electrocardiogram recordings. Comput Biol Med 39(1):88–96
4. Li Z et al (2021) A model for obstructive sleep apnea detection using a multi-layer feed-forward neural network based on electrocardiogram, pulse oxygen saturation, and body mass index. Sleep Breath 25(4):2065–2072
5. Atri R, Mohebbi M (2015) Obstructive sleep apnea detection using spectrum and bispectrum analysis of single-lead ECG signal. Physiol Meas 36(9):1963
6. Seisdedos VCR, Neto JE, Reyes EJM, Klautau A, de Oliveira RCL (2011) New approach for T-wave end detection on electrocardiogram: performance in noisy conditions. Biomed Eng Online 10(1):1–11

7. Indrawati AN et al (2022) Obstructive sleep apnea detection using frequency analysis of electrocardiographic RR interval and machine learning algorithms. J Biomed Phys Eng 12(6):627–636
8. Rajesh KNVPS, Dhuli R, Kumar TS (2021) Obstructive sleep apnea detection using discrete wavelet transform-based statistical features. Comput Biol Med 130:104199–104199
9. Gürüler H, Şahin M, Ferikoğlu A (2014) Feature selection on single-lead ECG for obstructive sleep apnea diagnosis. Turk J Electr Eng Comput Sci 22(2):465–478
10. Magnusdottir S, Hilmisson H (2017) Ambulatory screening tool for sleep apnea: analyzing a single-lead electrocardiogram signal (ECG). Sleep Breath 22:421–429
11. Hassan AR, Haque MA (2016) Computer-aided obstructive sleep apnea screening from single-lead electrocardiogram using statistical and spectral features and bootstrap aggregating. Biocybern Biomed Eng 36(1):256–266
12. Sharma H, Sharma KK (2020) Sleep apnea detection from ECG using variational mode decomposition. Biomed Phys Eng Express 6(1):015026
13. Penzel T et al (2000) The apnea-ECG database, pp 255–258; hu EAG (2000) Inside the fast Fourier transform black box: serial and parallel FFT algorithms. CRC Press, FL
14. Mahran AA, Hapsari RK, Nugroho H (2020) Penerapan Naive Bayes Gaussian Pada Klasifikasi Jenis Jamur Berdasarkan Ciri Statistik Orde Pertama. Netw Eng Res Oper 5(2):91–99
15. Al Shalabi LSZ, Kasasbeh B (2020) Data mining: a preprocessing engine. J Comput Sci 2006. 2(9)
16. Zhang Z (2016) Introduction to machine learning: k-nearest neighbors. Ann Transl Med 4(11):218–218
17. Byun H, Lee S (2003) A survey on pattern recognition applications of support vector machines. Int J Pattern Recognit Artif Intell 17(3):459–486
18. Ren J (2012) ANN vs. SVM: which one performs better in classification of MCCs in mammogram imaging. Knowl Based Syst 26:144–153
19. Vanage AM, Khade RH, Shinde DB (2012) Classifying five different arrhythmias by analyzing the ECG signals. IJCEM 15
20. Abidin Z (2012) Faktor Risiko Penyakit Jantung Koronerpada Pasien Rawat Inap Di cardiovascular care unit (CVCU) cardiac centre Rsup Dr. Wahidin Sudirohusodo Makassar periode Januari-Juli 2008
21. Lara F (1998) Artificial neural networks: an introduction. Instrum Dev 3(9)
22. Fukunaga K (2013) Introduction to statistical pattern recognition. Elsevier
23. Al Shalabi L, Shaaban Z, Kasasbeh B (2006) Data mining: a preprocessing engine. J Comput Sci 2(9):735–739
24. Zhu W, Zeng N, Wang NN (2010) Health care and life sciences sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS® implementations K&L consulting services, Inc., Fort Washington, PA Octagon Research Solutions, Wayne, pp 1–9
25. Mostafa SS et al (2020) Multi-objective hyperparameter optimization of convolutional neural network for obstructive sleep apnea detection. IEEE Access 8:129586–129599
26. Jain YK, Bhandare SK (2013) Min Max normalization based data perturbation method for privacy protection. Int J Comput Commun Technol 4(4)
27. Nguyen HD et al (2014) An online sleep apnea detection method based on recurrence quantification analysis. IEEE J Biomed Health Inform 18(4):1285–1293
28. Fatiah B et al (2020) Detection of apnea events from ECG segments using Fourier decomposition method. Biomed Signal Process Control 61:102005
29. Zhang X, Cheng J, Xu C, Lu H, Ma S (nd) LNEE 128—a dynamic batch sampling mode for SVM active learning in image retrieval
30. Geng Y, Tao D, Chen L, Li G (nd) LNEE 133—recognition of combined arm motions using support vector machine 2(133)
31. Teyakome J, Eiamkanitchat N (2015) A hybrid approach of neural network and level-2 fuzzy set. Lecture Notes Electr Eng 339:729–736

32. Bhardwaj A, Gupta A, Jain P, Rani A, Yadav J (2015) Classification of human emotions from EEG signals using SVM and LDA classifiers. In: International conference on signal processing and integrated networks (SPIN), Noida, India. IEEE, pp 180–185
33. Marcos JV et al (2009) Assessment of four statistical pattern recognition techniques to assist in obstructive sleep apnoea diagnosis from nocturnal oximetry. Med Eng Phys 31(8):971–978

The Power Use of Power Spectrum Density for Measures of Cognitive Performance Based on Electroencephalography: Systematic Literature Review

Rahmaniyah Dwi Astuti, Bambang Suhardi, Pringgo Widyo Laksono, Novie Susanto, and Ainun Rahmansyah Gaffar

Abstract Humans play more of a role as operators who carry out control functions, so cognitive abilities, especially those related to perception and decision-making, become very important. Cognitive performance can be seen in a person's ability to complete a cognitive activity. Electroencephalography is a tool for measuring cognitive performance through human brain activity wave signals. The main objective of this article is to systematically review Power Spectrum Density (PSD) feature extraction methods for cognitive performance measurement based on EEG. The methods used in this study are the PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analysis) method and Bibliometric analysis using VOSViewer. Data sources totaled 50 articles obtained from Scopus, Science Direct, and IOP Science for the 2013–2023 periods. The results of this article were obtained by observing keywords, density, article trends, feature extraction methods, and the application of EEG. The results showed that out of 50 articles that had been reviewed, 19 used PSD to measure cognitive performance, the Alpha frequency band is the most commonly used in measuring cognitive performance. The increasing use of PSD methods for the measurement of cognitive aspects (fatigue, workload, performance, mental) shows that future research directions can still be developed.

Keywords Power spectrum density · EEG · Cognitive performance · Systematic review

R. D. Astuti (✉) · B. Suhardi · P. W. Laksono · A. R. Gaffar
Department of Industrial Engineering, Faculty of Engineering, Sebelas Maret University, Surakarta 57126, Indonesia
e-mail: rahmaniyahdwis3@student.uns.ac.id

N. Susanto
Department of Industrial Engineering, Faculty of Engineering, Diponegoro University, Semarang 50275, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_12

1 Introduction

The human element is one of the most significant and foremost components of work in a working system. A good work system is inseparable from the workplace and the operational steps of tasks that must be carried out in a job. To create a work environment that is safe and fulfilling, the disciplines of human factors and ergonomics examine interactions between humans, machines, environments, and technology while taking human capabilities and limitations into account [1]. The development of industry today is very rapid, accompanied by modern and sophisticated technological advances. It is characterized by the change of production tools from manual to automatic. At the level of companies and industries, the development of technology and information and communication technology has changed the nature of work. Changes in work that were originally dominated by physical activity, into work or activities dominated by demands and mental or cognitive abilities. Humans play more of an operator role that carries out control functions, so cognitive abilities, especially those related to perception and decision-making, become very important. Physiological skills that only include motor skills and manual power can no longer be used as the only tool for analyzing human work performance. A complete and integrative approach is necessary for assessment in the design of work systems that combine cognitive demands in problem-solving, physical load in managing increasingly complex work systems, and interactions among people, work systems, and their surroundings. The focus should be shifted to looking at cognitive components in the age of digitalization and automation, where digital systems are becoming more common. There is increased physical and cognitive load as a result of our heavy reliance on technology.

Neuroscience is one of the fields of study of the nervous system in the human brain. Neuroscience also investigates biological, perceptual, and memory aspects of brain awareness and sensitivity. Neuroscience can make a connection between a cognitive process in the brain and the behavior that will be carried out by humans. This can be interpreted that, every command processed by the brain will activate important areas of the brain [2]. Researchers have been investigating brain mechanisms that enhance human interaction with technology in a naturalistic manner since the early 2000s with neuroergonomics, an interdisciplinary field that encompasses neuroscience, cognitive engineering, and human factors, aiming to enhance productivity and everyday experiences [3]. Neuroergonomics measures how the brain responds to different types of work. Neuroergonomics is crucial because it allows for flexible evaluation of brain function in realistic work environments [4, 5].

The application of different brain-imaging and psychophysiological methods is encouraged by neuroergonomics [3]. The more accurate the observation of brain wave activity, the easier it is to know a person's behavior and the more developed the science of neuroergonomics. One method that can be used to look at brain wave activity is the electroencephalography (EEG) [6]. The EEG is a method to record bioelectrical signals through a metal electrode on the scalp to determine the electrical activity of the brain [7]. So the EEG is a tool used to record or monitor brain activity. In order to

get maximum results, EEG is connected to BCI (Brain-Computer-Interface). BCIs are direct channel of communication pathways among the human brain and external devices [8]. The utilization of BCI with EEG is the process of extracting features and classifying EEG signals to connect the human mind with external equipment. BCIs objective is to convert brain activity into a command-to-control system; Measurable signals are used to transmit users' intentions. This data will be processing using special algorithms which can extract the required information from the EEG. Hence, to obtain the best signal ratio, the extraction step becomes important [9]. However, the EEG signal is challenging to understand because of its complexity. So the method is needed to extract useful features from EEG signals [10].

One method that is often used is Power Spectrum Density (PSD). The spectrum analysis of a signal involves separating it into several components according to its spectrum domain. In order to extract frequency features from a signal, the Fourier Transform (FT) must be used to describe its frequency components [11]. Different frequencies of brain activity can be isolated using the frequency feature. Typically, PSDs are calculated, followed by the extraction of the pertinent characteristics [12]. The purpose of this article is to conduct a systematic literature review of previous studies on the use of PSD as an extraction feature in EEG for cognitive performance. The results of this article are will help to assess the validity and reliability of the use of PSD as an extraction feature in EEG for cognitive performance.

2 Method

2.1 Review Standards

The systematic review of this article is based on guidelines in PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) and bibliometric analysis. PRISMA is a technical protocol that aids in ensuring accurate and significant review outcomes. The recommendations include 27 checklist items that guide researchers in gathering and accurately reporting information to raise the caliber of their work [13]. For investigating and analyzing vast amounts of scientific data, bibliometric analysis is a well-liked and exacting technique. It enables us to explore the subtleties of a field's evolutionary history while illuminating its frontiers [14]. Specifically targeted research questions and search strategies have been used in selecting articles to minimize the influence of expectations on the review.

2.2 Research Question

Study objectives guided the development of the following research questions (RQ):

- RQ1: What feature extraction is most commonly used when measuring cognitive performance using EEG?
- RQ2: What frequency bands are most often used during the EEG data extraction process using Power Spectrum Density (PSD)?
- RQ3: How is EEG data applied after extraction using the Power Spectral Density (PSD) feature?

2.3 Data Sources and Search Strategy

There are three database sources used in this study, namely Scopus, ScienceDirect and IOPscience. Content search is carried out on the three database sources, then filtered with the keywords "Power Spectral Density" OR "PSD" OR "Spectral Power Density" OR "Power Spectral Analysis" OR "Spectral Power Analysis" AND "EEG" OR "electroencephalogram" AND "Cognitive" OR "Cognitive Performance". Article search is limited by publication year, from 2013 to 2023. The last search conducted for this study was conducted on January 28, 2023. The subject area of the article is related to computer science, engineering, medicine, neuroscience, material science, and psychology.

2.4 Inclusion and Exclusion Criteria

In fulfilling eligibility, there are only a few articles selected with the following criteria (a) using English language article; (b) there is a quantitative EEG measurement using at least one frequency band (i.e., alpha, beta, gamma, delta, and theta) with spectral power density as feature extraction); (c) present up-to-date data; (d) Experiments involve only humans as subjects. The following characteristics of articles were excluded: (a) articles inaccessible and studies using overlapping data; (b) the study combined EEG with other brain electrical activity monitoring devices, such as MEG, fMRI, and fNIRS.

2.5 Study Selection

The first step is to find articles based on the research strategy, then delete articles that have the same title and filter based on article titles, following the flowchart sequence (Fig. 1). A screening process was followed to look for abstracts that met the inclusion

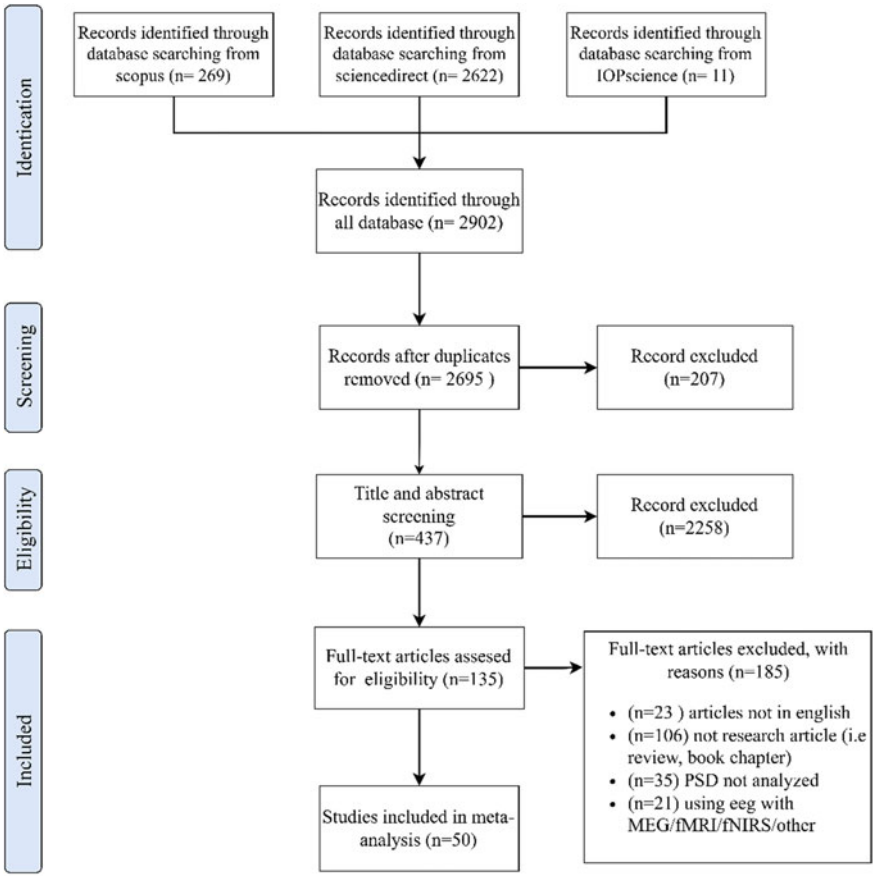


Fig. 1 The preferred reporting items for systematic review and meta-analyses (PRISMA) flow diagram

criteria, and those that did not were excluded. Full-text screening can be performed if the abstract doesn't provide sufficient information. Finally, following the review of full texts, we included studies which satisfied all inclusion criteria (n = 50).

2.6 Data Collection

Represented through the results of study characteristics. The data collected are (1) Author (2) Range age and gender of participants; (3) Frequency band; (4) Technique using; (5) Amount Channels; (6) Treatment; (7) Research Goal.

3 Results

3.1 Study Selection

Based on paper search keywords, 2902 articles were obtained. After deleting the same article total of 2695, screening the title and abstract with the appropriate articles was carried out with 437 articles were obtained. Based on the predetermined inclusion criteria, 185 articles did not meet the criteria, so as many articles to be reviewed (n = 50) were obtained. More clearly can be seen in Fig. 1. In addition, to present the article (n = 50) that has been obtained, the following is a publication journal on the article in Fig. 1.

3.2 Study Characteristic

The studies included, papers were published between 2013 and 2023 (mean: 2019 and 2019 median as well). A total of 1307 participants were obtained with various age ranges, the lowest age was 18 years, there are in articles [15–17]. While the oldest age was 64 years, there is article [18]. Some articles use mixed participants and some only use women as participants as in the article [19]. More details for characteristic data can be seen at the following link: https://bit.ly/AppendixTable of articles in Fig. 2.

3.3 Bibliometric Analysis

Keyword Analysis

Keyword network visualization shows the relationship of network visualization between topics adjusted to a predetermined keyword. Here are the network visualization keywords shown in Fig. 3. The results of keyword network visualization are obtained based on clusters, where there are a total of 103 items or keywords with 22 clusters. Here are the 5 clusters with the most keywords.

Density Analysis

Density visualization shows the strength and variety of research. The brighter the color in the image, it indicates higher occurrence, while if the color in the image is dim, it indicates lower occurrence. The following is the density visualization shown in Fig. 4.

The density visualization results show that based on articles (n = 50), “EEG”, “Power spectral density” and “cognitive” are the core keywords of the articles. In addition, several support keywords “noise”, “cognitive”, and “mental fatigue” that

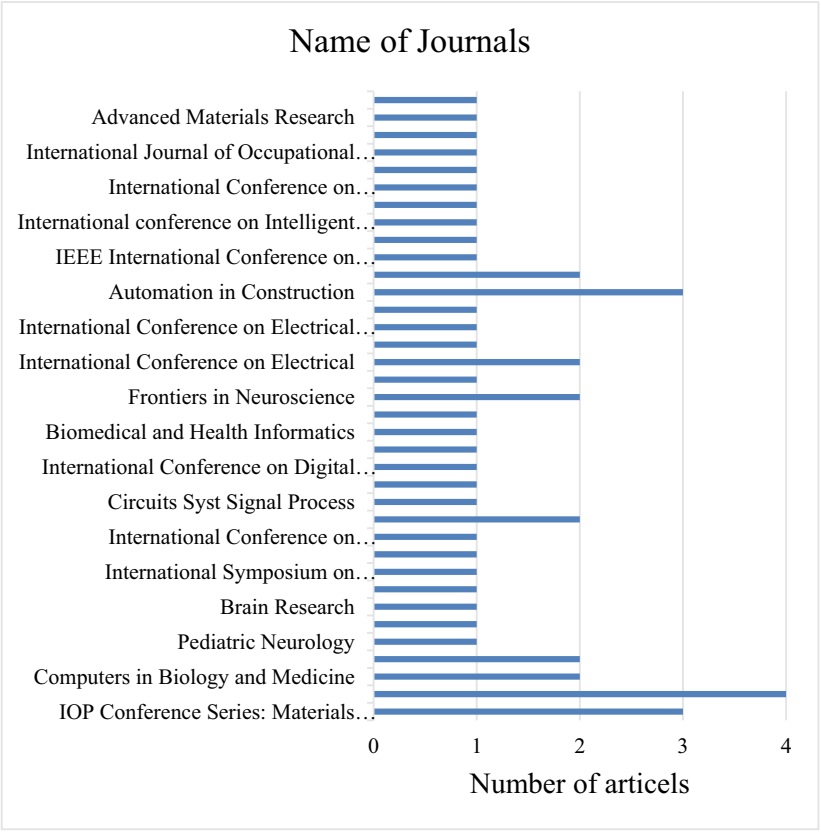


Fig. 2 Articles in journals published

indicate various aspects of measurement are widely used in conjunction with "EEG and Power Spectral Density".

Overlay Analysis

Overlay visualization shows the range of years published articles (n = 50) based on the keyword articles. Here is the overlay visualization shown in Fig. 5.

The overlay visualization results show that the most keyword articles (n = 50) were in 2019. Similarly, core keywords were published in 2019 while support keywords "noise" and "cognitive" were published in 2017 and 2018.

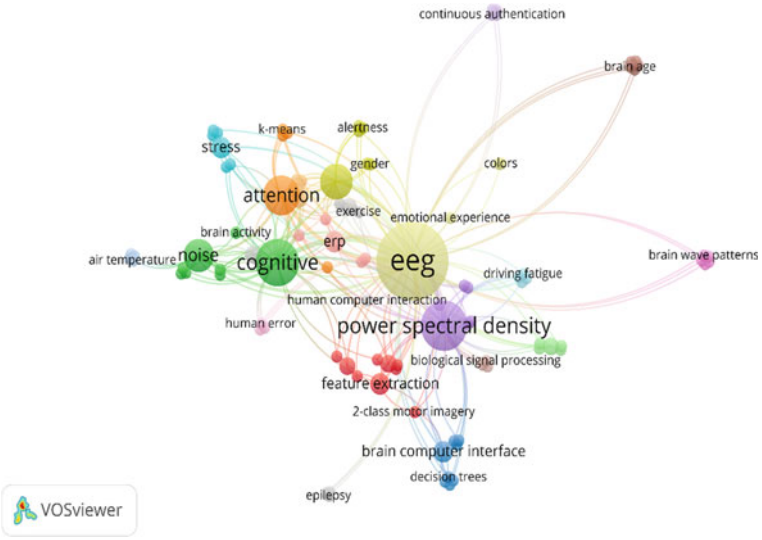


Fig. 3 Keyword network visualization

Table 1 5 clusters with the most keywords

Cluster	Keyword	Total
First cluster (red)	2-class motor imagery, channel selection, classification, concentration, feature extraction, fft, perceived stess, short-time fourier transform, svm	9
Second cluster (dark green)	Annoyance, beta signal, brain activity, cognitive, muse headband, noise, stroop test, task difficulty	8
Third cluster (dark blue)	Brain computer interface, decision trees, few channel, motor imagery, psychology experiment, random forest, situational awareness, time window optimazation	8
Fourth cluster (yellow)	Alertness, drowsiness, (eeg, mental fatigue, alartness), gender, mental fatigue, psychoacoustic	6
Five cluster (purple)	Berger effect, depth of anesthesia, hypoxic-ischemic encephalopathy, ordinal representation, physiologic biomarker	6

3.4 Content Analysis

Feature Extraction

Based on the results of the paper that have been obtained (n = 50), RQ1 can be classified as shown in Fig. 6.

On Fig. 6, there are 43% or 26 articles, using PSD as feature extraction, including [10, 15–39] and so on. There is also an article that uses PSD and ERP (Event Related

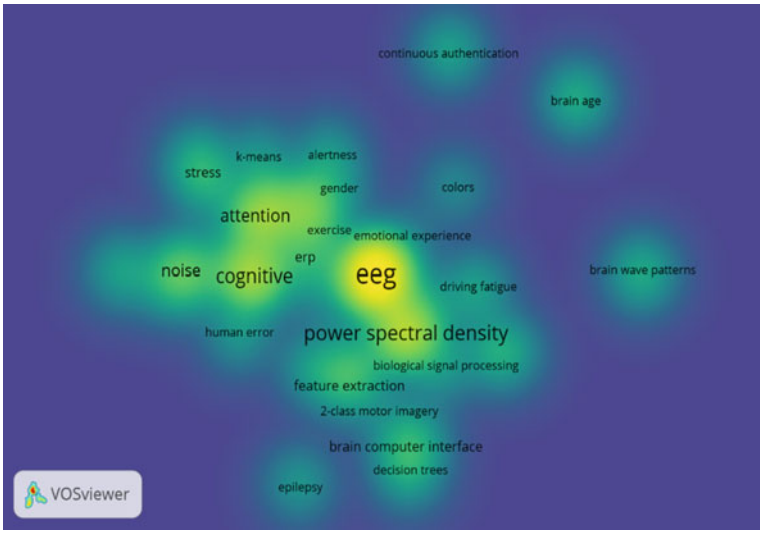


Fig. 4 Density visualization

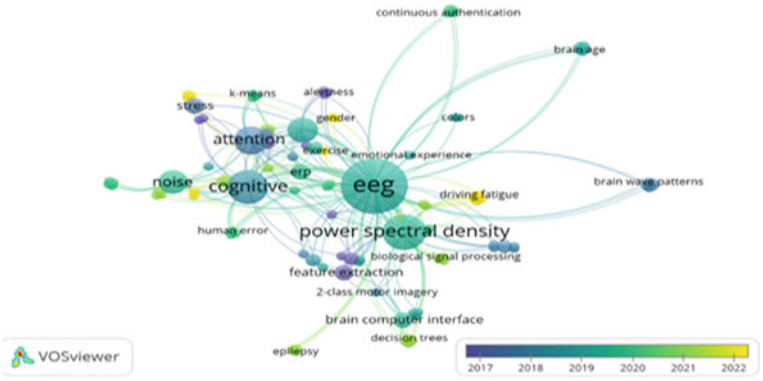


Fig. 5 Overlay visualization

Potentials) for feature extraction, namely [24]. However, there are 20%, or 12 articles, that explain welch methods, including [26, 36, 40–49]. Some articles explain how to convert the obtained EEG data to the frequency (spectral) domain so that such fluctuations can be extracted, namely 7% or 4 articles using STFT (Short Time Fourier Transform [16, 31, 49, 50] and 23% or 14 articles using FFT (Fast Fourier Transform), including [6, 17, 28, 51–61]. In addition, 10%, or 6 articles, use wavelet, DWT (Discrete Wavelet Transform), power band, HHT (Hilbert-Huang Transform), and WT (Wavelet Transform). Including [17, 27, 30, 38, 50, 62].

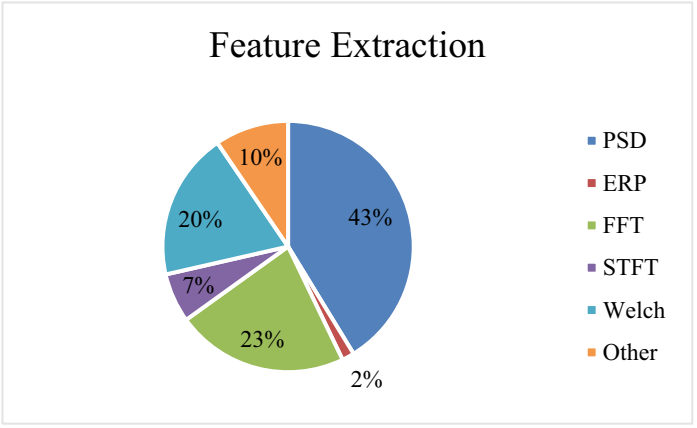


Fig. 6 Percentage of technique used PSD

Frequency Band

Based on the results of the paper that has been melted (n = 50), RQ2 can be classified as follows:

On Fig. 7, 32% of articles use frequency band alpha (8–12 Hz) related activity for creativity, relaxation, and visualization. 24% use frequency band beta (12–27 Hz) related activity for beware and concentration. 9% use frequency band delta (0.5–3 Hz) related activity for cure and sleep very well, and 23% use frequency band theta (3–8 Hz) related activity for deep relaxation, meditation, and improved memory. In addition, some articles only use one frequency band, namely article [22] using delta, article [23] using alpha, article [26] using delta, article [37] using alpha, article [39] using theta, article [42] using delta, article [47] using gamma, article [48] using beta, article [51] using alpha, and article [55] using alpha.

Application of EEG

Based on the results of the paper that have been obtained (n = 50), RQ3 can be classified as follows:

Besides Fig. 8, each article has a research goal that can be classified as follows; Articles classified as mental classification (workload, fatigue) include article [20] has a research goal for the detection of driving fatigue. Article [22] has a research goal for emotion recognition. Article [25] has a research goal for classifying pleasure and mental displeasure. Article [31] has a research goal of indicating different levels of mental factors. Article [41] has a research goal stress analysis. Article [49] has a research goal of calculating an operator's mental effort based on cognitive task difficulty levels. Article [54] has a research goal of examining the touching effect of frustrating emotion in the nursing setting. The articles classified as cognitive (fatigue, workload, performance), among others [19], have a research goal for measures of cognitive performance. Article [23] has a research goal for the estimation of cognitive

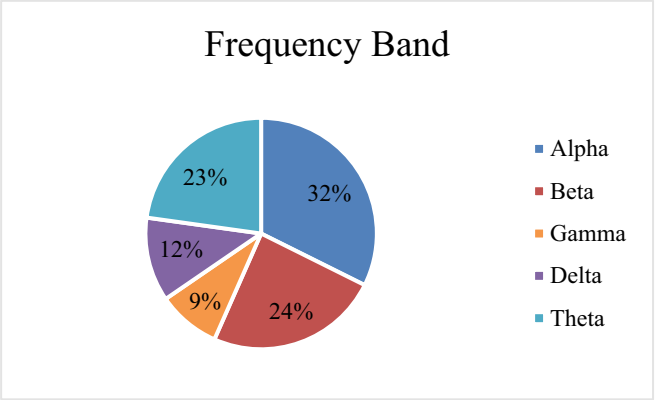


Fig. 7 Percentage of frequency band

states. Article [36] has a research goal for calculating indicators of cognitive fatigue. Article [38] has a research goal combining noise and temperature to evaluate their effects. Article [39] has a research goal of measuring in chemical process control rooms, human operators have a high cognitive workload. As for Articles classified as attention or concentration, among others, article [17] has a research goal for measures of human concentration. Article [44] has a research goal to improve the recognition performance of different odor concentrations. Article [52] has a research goal for attentional state recognition. As for articles classified as medicine, among others, article [42] has a research goal for classifying the hypoxic-ischemic. Article [51] has the research goal of a new score based on Propofol dosages will be computed to compare the incidence of Burst-suppressions (BS) with age. The articles classified as BCI classification, namely article [48], determine the best time period for thinking tasks for each subject by conducting research. Article [50] has a research goal of proposing an effective feature extraction method, WDPSD, for a 2-class motor imagery-based BCI system. The articles classified as Situation Awareness, among others, article [26] has a research goal of a comparison of measurements in a room without and with a Faraday cage. Article [45] has a research goal for the measurement of resting state. Article [55] has a research goal for observe how the brain perceives the world and how calm it is. Articles classified as Authentication include, article [46] has a research goal of examining the feasibility of using fictitious or non-fictitious tasks for human authentication. Based on the classification that has been done ($n = 50$), the application of PSD as feature extraction is mostly carried out in the mental classification (workload, fatigue) by as much as 34% and the least in the BCI classification by as much as 4%.

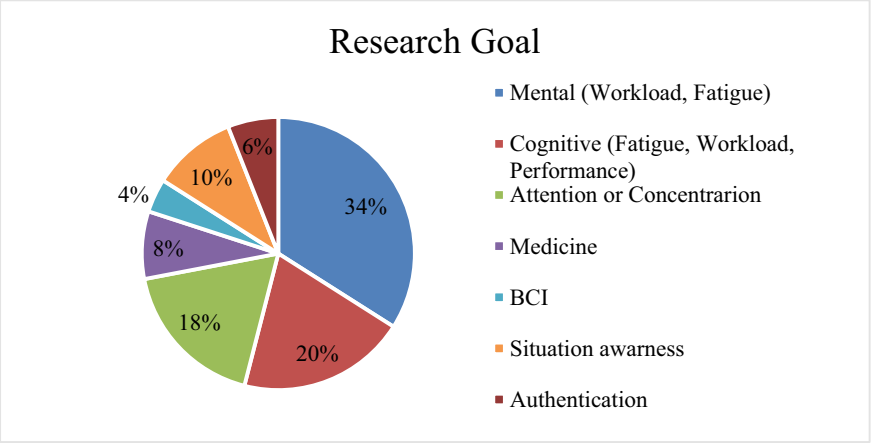


Fig. 8 Percentage of research goal

4 Discussion

Analyzing cognitive performance using EEG with power spectral density as a feature extraction technique is a topic of previous research in this review article. The characteristics of articles (n = 50) can be seen in the https://bit.ly/AppendixTableofPaper. Among the 50 articles, there is a research goal that does not include the measurement of cognitive performance. The total number of articles that focus on measuring cognitive performance amounted to 19 articles. 9 articles focus on measuring attention and concentration, 4 of which focus on the impact of noise on cognitive performance but use different cognitive tasks. While the other 5 articles categorize the level of attention and concentration of participants (the impact of the stimulus given on the attention and concentration of participants). Of the 9 articles, most use the alpha and beta frequency bands, this is because the alpha frequency band can measure the level of relaxation while the beta frequency band can measure the level of concentration and attention. 10 articles focus on the cognitive measurement of fatigue, workload, and performance. 2 papers focused on cognitive fatigue and cognitive performance with different stimuli for participants. The process of processing the signal of each article by transforming the time domain signal into the frequency domain so that it can produce power spectral density as feature extraction, the transformation tends to use a fast Fourier transform. There are also similarities in each article that uses an experimental design with the highest number of participants reaching 120 participants, both male, and male, although each article that uses mixed participants has not classified gender against EEG measurements carried out. A systematic literature review that focuses on power spectral density to measure cognitive performance has not been widely done, In an article [63] measuring cognitive workload using power spectral, the article uses 45 effect sizes from 24 studies. 17 effect sizes use frequency band alpha, 16 effect sizes use frequency band theta, and 12 effect sizes use frequency

band beta. This is in line with this article, where frequency bands alpha and beta are most often used when extracting EEG data.

Review this article's direction or recommendations for further research. Based on 50 articles that have been reviewed, cognitive performance measurement using power spectral density as EEG feature extraction is included in the neuroergonomics branch and has a fairly broad scope. PSD as feature extraction itself can be used in many fields of research such as psychology, materials science, and engineering. Research on cognitive performance measurement can still be developed further, as in previous studies by measuring the level of attention, concentration, and impact noise on performance. However, classifying gender against performance measurement has not been widely done by researchers. So considering gender classification can improve the quality of research because it has more specific results.

Based on the methodology of this article, there are several limitations that can affect the results of future research, systematic literature review does not cover all relevant scopes to the use of power spectral density. This article also only looks at the trends of the article 10 years ago, when compared to article [63], which looks at the trends of the article 35 years ago, and article [64], which looks at trends 18 years ago, with studies included in the literature review as many as 143. So articles need more study, including trend articles, to get more accurate results. However, the results of this article can be used as a reference for researchers to use the power spectral density feature to measure cognitive aspects.

5 Conclusion

Based on a articles (n = 50) that have been carried out using PRISMA and bibliometric methods to systematically review Power Spectrum Density (PSD) feature extraction methods for cognitive performance measurement based on EEG, it was obtained what feature extraction results are most often used when measuring cognitive performance using EEG, frequency bands that are most often used when extracting EEG data, and applying EEG using PSD as an extraction feature. PSD is feature extraction, there are 19 used PSD to measure cognitive performance. Alpha signal is the most widely used frequency, there are 32% of articles use frequency band alpha. The article aims to measure mental and cognitive. In addition, this review paper can be used as a reference for the development of EEG or Power Spectral Density especially in the field of neuroergonomics.

References

1. Karwowski W (2005) Ergonomics and human factors: the paradigms for science, engineering, design, technology and management of human-compatible systems. Ergonomics 48(5):436–463. https://doi.org/10.1080/00140130400029167
2. Hengki W (2018) Pendidikan Neurosains Dan Implikasinya Dalam Pendidikan Masa Kini. Pendidik Dasar 2(March):1–19
3. Dehais F, Ayaz H (2018) Progress and direction in neuroergonomics, no 2017. Elsevier
4. Mehta RK, Parasuraman R (2013) Neuroergonomics: a review of applications to physical and cognitive work. Front Hum Neurosci 7(Dec):1–10. https://doi.org/10.3389/fnhum.2013.00889
5. Parasuraman R, Wilson GF (2008) Putting the brain to work: neuroergonomics past, present, and future. Hum Factors 50(3):468–474. https://doi.org/10.1518/001872008X288349
6. Jafari MJ, Khosrowabadi R, Khodakarim S, Mohammadian F (2019) The effect of noise exposure on cognitive performance and brain activity patterns. Open Access Maced J Med Sci 7(17):2924–2931. https://doi.org/10.3889/oamjms.2019.742
7. Cascino GD (1991) Current practice of clinical electroencephalography, 2nd ed, vol 41, no 3
8. Trad D, Al-Ani T, Monacelli E, Delaplace S, Jemni M (2011) Feature extraction based on empirical mode decomposition and band power approaches for motor imagery tasks classification. In: Proceedings of the IADIS international conference on interfaces human-computer interaction 2011, Part IADIS multi conference on computer science and information systems 2011, MCCSIS 2011, no. April 2015, pp 185–192
9. Agustina Garcés M, Orosco LL (2018) EEG signal processing in brain-computer interface, 2nd ed. Elsevier B.V.
10. Kumar N, Kumar J (2016) Measurement of cognitive load in HCI systems using EEG power spectrum: an experimental study. Procedia Comput Sci 84:70–78. https://doi.org/10.1016/j.procs.2016.04.068
11. Akrami A, Solhjoo S, Motie-Nasrabadi A, Hashemi-Golpayegani MR (2005) EEG-based mental task classification: linear and nonlinear classification of movement imagery. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society. VOLS, vol 7, pp 4626–4629. https://doi.org/10.1109/iembs.2005.1615501
12. Varsavsky A, Mareels I, Cook M (2011) EPILEPTIC and the EEG EPILEPTIC and the EEG
13. Moher D et al (2010) CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 340. https://doi.org/10.1136/bmj.c869
14. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM (2021) How to conduct a bibliometric analysis: an overview and guidelines. J Bus Res 133(April):285–296. https://doi.org/10.1016/j.jbusres.2021.04.070
15. Kästle JL, Anvari B, Krol J, Wurdemann HA (2021) Correlation between situational awareness and EEG signals. Neurocomputing 432:70–79. https://doi.org/10.1016/j.neucom.2020.12.026
16. Mohamed Z, El Halaby M, Said T, Shawky D, Badawi A (2018) Characterizing focused attention and working memory using EEG. Sensors (Switzerland) 18(11):1–21. https://doi.org/10.3390/s18113743
17. Purnamasari PD, Junika TW (2019) Frequency-based EEG human concentration detection system methods with SVM classification. In: Proceedings: CYBERNETICSCOM 2019 - 2019 IEEE international conference on cybernetics and computational intelligence: towards a smart and human-centered cyber world, pp 29–34. https://doi.org/10.1109/CYBERNETICSCOM.2019.8875677
18. Pieper K et al (2021) Working with environmental noise and noise-cancelation: a workload assessment With EEG and subjective measures. Front Neurosci 15(November):1–13. https://doi.org/10.3389/fnins.2021.771533
19. Foy JG, Foy MR (2020) Dynamic changes in EEG power spectral densities during NIH-toolbox Flanker, dimensional change card sort test and episodic memory tests in young adults. Front Hum Neurosci 14(May):1–10. https://doi.org/10.3389/fnhum.2020.00158

20. Zheng Y, Ma Y, Cammon J, Zhang S, Zhang J, Zhang Y (2022) A new feature selection approach for driving fatigue EEG detection with a modified machine learning algorithm. Comput Biol Med 147(May):105718. https://doi.org/10.1016/j.compbiomed.2022.105718
21. Iqbal MU, Shahab MA, Choudhary M, Srinivasan B, Srinivasan R (2021) Electroencephalography (EEG) based cognitive measures for evaluating the effectiveness of operator training. Process Saf Environ Prot 150:51–67. https://doi.org/10.1016/j.psep.2021.03.050
22. Rahman MA, Anjum A, Milu MMH, Khanam F, Uddin MS, Mollah MN (2021) Emotion recognition from EEG-based relative power spectral topography using convolutional neural network. Array 11(June):100072. https://doi.org/10.1016/j.array.2021.100072
23. Islam M, Ahmed T, Yusuf MSU, Ahmad M (2014) Channel selection and feature extraction for cognitive state estimation with the variation of brain signal. In: 2013 international conference on electrical information and communication technology EICT 2013. https://doi.org/10.1109/EICT.2014.6777860
24. Gentili RJ et al (2018) Combined assessment of attentional reserve and cognitive-motor effort under various levels of challenge with a dry EEG system. Psychophysiology 55(6):1–17. https://doi.org/10.1111/psyp.13059
25. Wong RZ, Choo YH, Muda AK (2020) Task sensitivity in continuous electroencephalogram person authentication. Int J Adv Comput Sci Appl 11(2):552–558. https://doi.org/10.14569/ijacsa.2020.0110270
26. Suwandi GRF, Khotimah SN, Suprijadi (2022) Electroencephalography signal power spectral density from measurements in room with and without faraday cage: a comparative study. J Phys Conf Ser 2243(1). https://doi.org/10.1088/1742-6596/2243/1/012002
27. Tseng LH, Cheng MT, Chen ST, Hwang JF, Chen CJ, Chou CY (2013) An EEG investigation of the impact of noise on attention. Adv Mater Res 779:1731–1736. https://doi.org/10.4028/www.scientific.net/AMR.779-780.1731
28. Arsalan A, Majid M, Butt AR, Anwar SM (2019) Classification of perceived mental stress using a commercially available EEG headband. IEEE J Biomed Heal Inform 23(6):2257–2264. https://doi.org/10.1109/JBHI.2019.2926407
29. Abbasi AM, Motamedzade M, Aliabadi M, Golmohammadi R, Tapak L (2018) Study of the physiological and mental health effects caused by exposure to low-frequency noise in a simulated control room. Build Acoust 25(3):233–248. https://doi.org/10.1177/1351010X18779518
30. Ke J, Du J, Luo J (2021) The effect of noise content and level on cognitive performance measured by electroencephalography (EEG). Autom Constr 130. https://doi.org/10.1016/j.autcon.2021.103836
31. Dasari D, Shou G, Ding L (2017) ICA-derived EEG correlates to mental fatigue, effort, and workload in a realistically simulated air traffic control task. Front Neurosci 11(May). https://doi.org/10.3389/fnins.2017.00297
32. Trejo LJ, Kubitz K, Rosipal R, Kochavi RL, Montgomery LD (2015) EEG-based estimation and classification of mental fatigue. Psychology 06(05):572–589. https://doi.org/10.4236/psych.2015.65055
33. Ong ZY, Saidatul A, Ibrahim Z (2018) Power spectral density analysis for human EEG-based biometric identification. In: 2018 international conference on computational approach in smart systems design and applications, ICASSDA 2018, no. August, pp 1–6. https://doi.org/10.1109/ICASSDA.2018.8477604
34. Wang S et al (2022) Modulating break types induces divergent low band EEG processes during post-break improvement: a power spectral analysis. Front Hum Neurosci 16. https://doi.org/10.3389/fnhum.2022.960286
35. Jebelli H, Hwang S, Lee SH (2018) EEG-based workers' stress recognition at construction sites. Autom Constr 93(January):315–324. https://doi.org/10.1016/j.autcon.2018.05.027
36. Aziezah F, Harke Pratama S, Yulianti, Wahidah S, Haryanto F, Suprijadi (2020) Characterization of individual alpha frequency of EEG signals as an indicator of cognitive fatigue. J Phys Conf Ser 1505(1). https://doi.org/10.1088/1742-6596/1505/1/012068

37. Chen Z, Lin L (2019) Emotional experience evaluation method of interaction task based on EEG technology. IOP Conf Ser Mater Sci Eng 573(1). https://doi.org/10.1088/1757-899X/573/1/012022
38. Abbasi AM, Motamedzade M, Aliabadi M, Golmohammadi R, Tapak L (2020) Combined effects of noise and air temperature on human neurophysiological responses in a simulated indoor environment. Appl Ergon 88(June):103189. https://doi.org/10.1016/j.apergo.2020.103189
39. Iqbal MU, Srinivasan B, Srinivasan R (2020) Dynamic assessment of control room operator's cognitive workload using electroencephalography (EEG). Comput Chem Eng 141:106726. https://doi.org/10.1016/j.compchemeng.2020.106726
40. Ameera A, Saidatul A, Ibrahim Z (2019) Analysis of EEG spectrum bands using power spectral density for pleasure and displeasure state. IOP Conf Ser Mater Sci Eng 557(1). https://doi.org/10.1088/1757-899X/557/1/012030
41. Roy S, Islam M, Yusuf MSU, Jahan N (2022) EEG based stress analysis using rhythm specific spectral feature for video game play. Comput Biol Med 148(May):105849. https://doi.org/10.1016/j.compbiomed.2022.105849
42. Kota S, Jasti K, Liu Y, Liu H, Zhang R, Chalak L (2021) EEG spectral power: a proposed physiological biomarker to classify the hypoxic-ischemic encephalopathy severity in real time. Pediatr Neurol 122:7–14. https://doi.org/10.1016/j.pediatrneurol.2021.06.001
43. Lanzone J et al (2021) The effect of Perampanel on EEG spectral power and connectivity in patients with focal epilepsy. Clin Neurophysiol 132(9):2176–2183. https://doi.org/10.1016/j.clinph.2021.05.026
44. Hou HR, Meng QH (2021) A double-square-based electrode sequence learning method for odor concentration identification using EEG signals. IEEE Trans Instrum Meas 70. https://doi.org/10.1109/TIM.2021.3090177
45. Dirik HB, Darendeli A, Ertan H (2022) The new wireless EEG device Mentalab explore is a valid and reliable system for the measurement of resting state EEG spectral features. Brain Res 1798(November):148164. https://doi.org/10.1016/j.brainres.2022.148164
46. Zhi Chin T, Saidatul A, Ibrahim Z (2019) Exploring EEG based authentication for imaginary and non-imaginary tasks using power spectral density method. IOP Conf Ser Mater Sci Eng 557(1). https://doi.org/10.1088/1757-899X/557/1/012031
47. Thomas KP, Vinod AP (2018) EEG-based biometric authentication using gamma band power during rest state. Circuits Syst Signal Process 37(1):277–289. https://doi.org/10.1007/s00034-017-0551-4
48. Yang B, Hu C, Wang J, Li B, Wang W (2019) Research on motor imagery EEG modeling based on window optimization and a few channels PSD. In: International conference on digital signal processing, DSP, vol 2018-November, 2019. https://doi.org/10.1109/ICDSP.2018.8631698
49. Zammouri A et al (2017) Brain waves-based index for workload estimation and mental effort engagement recognition. J Phys Conf Ser 904(1). https://doi.org/10.1088/1742-6596/904/1/012008
50. Kim C, Sun J, Liu D, Wang Q, Paek S (2018) An effective feature extraction method by power spectral density of EEG signal for 2-class motor imagery-based BCI. Med Biol Eng Comput 56(9):1645–1658. https://doi.org/10.1007/s11517-017-1761-4
51. Touchard C et al (2019) EEG power spectral density under Propofol and its association with burst suppression, a marker of cerebral fragility. Clin Neurophysiol 130(8):1311–1319. https://doi.org/10.1016/j.clinph.2019.05.014
52. Tang Y, Huang H (2020) An EEG-based brain-computer interface for attention state recognition. In: 2020 international symposium on autonomous systems ISAS 2020, pp 100–104. https://doi.org/10.1109/ISAS49493.2020.9378873
53. Lee H, Kim Y, Park C, Subjects A (2018) Classification of human attention to media lecture. In: International conference on information networking, pp 914–916
54. Muramatsu T, Washizawa Y, Hiyoshi K (2020) EEG analysis of nursing touch for frustrating work. In: LifeTech 2020—2020 IEEE 2nd global conference on life sciences and technologies, LifeTech, pp 67–71. https://doi.org/10.1109/LifeTech48969.2020.1570617576

55. Azwar SHNS, Amin MKM, Islam AKMM, Mikami O (2019) Electroencephalogram (EEG) studies on human perception in colours. J Adv Manuf Technol 13(Special Issue 1):163–174
56. Choi HI, Noh GJ, Shin HC (2020) Measuring the depth of anesthesia using ordinal power spectral density of electroencephalogram. IEEE Access 8:50431–50438. https://doi.org/10.1109/ACCESS.2020.2980370
57. Jun G, Smitha KG (2017) EEG based stress level identification. In: 2016 IEEE international conference on systems, man, and cybernetics SMC 2016, pp 3270–3274. https://doi.org/10.1109/SMC.2016.7844738
58. Putman P, Verkuil B, Arias-Garcia E, Pantazi I, Van Schie C (2014) EEG theta/beta ratio as a potential biomarker for attentional control and resilience against deleterious effects of stress on attention. Cogn Affect Behav Neurosci 14(2):782–791. https://doi.org/10.3758/s13415-013-0238-7
59. Sadeghian M, Mohammadi Z, Mousavi SM (2021) Investigation of electroencephalography variations of mental workload in the exposure of the psychoacoustic in both male and female groups. Cogn Neurodyn, November 2021. https://doi.org/10.1007/s11571-021-09737-3
60. Sadeghian M et al (2021) Effect of tonal noise and task difficulty on electroencephalography and cognitive performance. Int J Occup Saf Ergon 1–9. https://doi.org/10.1080/10803548.2021.1901432
61. Ke J, Zhang M, Luo X, Chen J (2021) Monitoring distraction of construction workers caused by noise using a wearable electroencephalography (EEG) device. Autom Constr 125(February):103598. https://doi.org/10.1016/j.autcon.2021.103598
62. Borghini G et al (2014) Analysis of neurophysiological signals for the training and mental workload assessment of ATCos. In: SIDs 2014 – Proceedings of SESAR innovation days
63. Chikhi S, Matton N, Blanchet S (2022) EEG power spectral measures of cognitive workload: a meta-analysis. Psychophysiology 59(6):1–24. https://doi.org/10.1111/psyp.14009
64. Ismail LE, Karwowski W (2020) Applications of EEG indices for the quantification of human cognitive performance: a systematic review and bibliometric analysis. 15(12)

Enhancing Rice Leaf Disease Classification: A Combined Algorithm Approach for Improved Accuracy and Robustness

Apri Junaidi, Diao Qi, Chan Weng Howe, and Siti Zaiton Mohd Hashim

Abstract This research addresses the problem of improving image classification accuracy, given the importance of classification accuracy in applications such as disease diagnosis and object recognition. This research aims to explore various deep-learning architectures and ensemble methods to improve classification accuracy effectively. We use a comprehensive approach for methodology, evaluating training data from various architectures using 17 transfer learning models. Our method incorporates models such as EfficientNetB0, AlexNet, and MobileNetV2, utilizing their unique strengths. We also designed a two-stage model that progressively combines architectures between low-accuracy and two high-accuracy models to create a more accurate classifier. The dataset contains 5932 images of rice leaf diseases distributed into four classes: Bacterial Leaf Blight, Blast, Brown Spot, and Tungro. The outstanding results showed a substantial increase in accuracy from 0.35 to 0.97 in the ensemble model. This significant improvement underscores the potential of integrating different architectures to utilize complementary features, ultimately improving classification accuracy. This research provides insights into image classification and offers practical solutions to improve accuracy in various domains. As for the implications, this study shows the promise of a blended approach to deep learning architecture in significantly improving image classification performance in various domains.

Keywords Ensemble method · Combine algorithm · Two-stage model development · Image classification · Transfer learning

A. Junaidi (✉) · D. Qi · C. W. Howe · S. Z. M. Hashim
Faculty of Computing, Universiti Teknologi Malaysia, UTM Johor Bahru, Johor 81310, Malaysia
e-mail: junaidi20@graduate.utm.my

D. Qi
Faculty of Artificial Intelligence, Zhejiang Dongfang Polytechnic, Wenzhou 32500, China

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_13

1 Introduction

Rice (Oryza sativa) cultivation plays an important role in global food security, providing a staple food source for most of the world's population [1]. However, the yield and quality of the rice crop are constantly threatened by various diseases that attack the plant leaves. Bacterial Leaf Blight, Blast, Brown Spot, and Tungro are common diseases that attack rice leaves [2]. Techniques are needed to be able to classify the types of rice leaf disease types [3]. From several techniques used, such as transfer learning, Convolution Neural Networks and other classification techniques [4, 5], While previous studies seem to have only used a few pre-trained models that did have a good model structure, they did not include some models that did not achieve high accuracy values.

Sethy et al. [6] this study discusses the use of the deep convolutional neural network (CNN) to identify four types of rice leaf diseases using field images. The advantages of this research are training eleven CNN models with a transfer learning approach, and deep features plus support vector machine (SVM) are evaluated, with the results of deep features plus SVM having better performance in classification, but the disadvantage of this research is that this research does not train pre-trained models that have low accuracy, so it is possible to do ensemble learning for models with low accuracy. The study [7] evaluates how well a convolutional neural network (CNN) design classifies photos of diseases affecting rice leaves. In this study, investigate the accuracy of different CNN structures in classifying a dataset of 5932 photos divided into four disease groups. According to the study's empirical results, the InceptionV3 and InceptionResnetV2 architectures achieved 100% accuracy and the highest precision, followed by ResNet50 and DenseNet201 at 99.83%, MobileNet at 99.33%, and EfficientNetB3 at 90.14%, Advantage of this research is evaluating the performance of different CNN architectures, which can help in selecting the best architecture for the classification of rice leaf disease images. The disadvantage for this research is limited to the accuracy of the CNN architectures may vary for other disease classes. The study does not consider the impact of environmental factors on the accuracy of the CNN architectures. Krishnamoorthy et al. [8] proposes the use of artificial neural networks with transfer learning to predict rice leaf diseases. This study experimented with pre-trained models and simple CNNs with training and testing sets consisting of 4000 and 300 examples respectively. The InceptionResNetV2 model was used with a transfer learning approach to recognize diseases in rice leaf images, achieving a good accuracy of 95.67%, this research have advantage use of deep neural networks with transfer learning approach can greatly assist in automatic identification of infectious organisms in the rice plant leaves, which can save time and resources for farmers, the other hand disadvantage of this research is the researcher did not compare the proposed model with other state-of-the-art models for rice leaf disease detection, which makes it difficult to assess the performance of the proposed model in comparison to other models. In this research the researcher [9] investigates the application of transfer learning within intricate convolutional neural networks (CNNs); the tests encompass the division of the dataset into varying ratios for training and testing sets.

The suggested model has the capability to categorize rice diseases, achieving a classification accuracy of 91.37% when utilizing an 80%–20% training–testing split. The advantage in this research is that the method has the potential to be expanded for the creation of a practical agricultural field system dedicated to identifying diseases in rice plants, disadvantages is the method might not exhibit effectiveness in detecting and categorizing every variety of rice plant diseases. The research [10] suggests an application of deep learning through Convolutional Neural Network (CNN) to automatically classify three specific types of rice leaf diseases: bacterial blight, blast, and brown mark. The CNN model demonstrates a capability to differentiate between healthy and diseased leaves with an accuracy of 94%, while achieving a classification accuracy of 78.44% for various types of diseased rice leaves. This research has advantage is the CNN model has the capacity to acquire essential features directly from raw images, thus obviating the necessity for manual identification of relevant attributes to differentiate between various types of crop diseases. On the other hand, the CNN model's accuracy in classifying various types of diseased rice leaves stands at 78.44%, indicating that there is still potential for enhancement. Bhowmik et al. [11] introduces an ensemble learning approach using transfer learning and LightGBM to detect rice leaf diseases. This method managed to achieve a high accuracy rate of 0.96, but it also has some limitations that need to be considered in further development, such as testing LightGBM with other transfer learning models, so that the test will appear balanced, and it will be known which transfer learning model is better when combined with LightGBM. The development of image classification models commonly involves the use of three techniques: Convolutional Neural Networks (CNN) [12], Transfer Learning, and Ensemble Learning. CNN is employed for the detection of visual patterns in images through the use of complex layers. Transfer Learning leverages pre-existing knowledge from previously trained models to improve the performance of the classification model. Ensemble learning, on the other hand, combines multiple classification models to achieve more precise results by aggregating predictions from each model [13].

Therefore, this study aims to classify rice leaf diseases. This study performs classification with 17 transfer learning models, of these 17 models not all models produce accuracy values above 90%. For the model that has the lowest accuracy with a value of 38%, model training is carried out with ensemble learning. The ensemble learning stage is carried out in two stages, the first stage performs ensemble learning between AlexNet and EfficientNetB0 models, resulting in an accuracy of 43%. The second stage between AlexNet, MobileNetV2 and EfficientNetB0, this second stage ensemble learning produces 95% accuracy. Therefore, the method of combining models that have low accuracy with several models that have high accuracy values, based on this research, can produce high accuracy values as well.

2 Materials and Method

The rice leaf disease classification model conducted in this study can be seen in Fig. 1. that explains the proposed method in this study, there are three phases that are carried out starting from dataset collection, then entering the first phase which is to enhance data performance by doing data preprocessing.

Then stepping on the second phase to aim for the best accuracy, by going through several steps namely create base model configuration, compile model and fit model. in this section is known as transfer learning and ensemble learning, the difference is the formation of ensemble models that combine transfer learning models that have low accuracy values with two models that have high accuracy values, where researchers use seventeen pre-trained models, as part of a new model architecture for the problem of classifying rice leaf disease types.

The last phase gets the accuracy of seventeen models that have been trained with the dataset of rice leaf disease types, the accuracy value produced by the model shows how precise the model is in classifying rice leaf diseases.

2.1 Dataset

Based on previous research, this research follows the same dataset used by previous researchers [6], the dataset can be obtained at the url: https://data.mendeley.com/datasets/fwcj7stb8r/1, The data collection includes 5932 number images and four different types of diseases that affect rice leaves, Fig. 2 represents the form of the

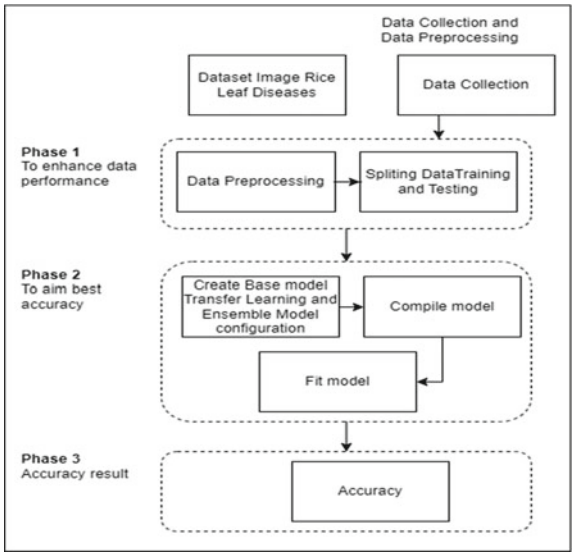


Fig. 1 Proposed method for rice leaf diseases classification



(a)Tungro (b) Brown spot (c) Blast (d) Bacterial blight

Fig. 2 Example image of rice leaf disease dataset by class

image in the dataset consisting of four classes, namely, (a) Tungro, (b) Brown spot, (c) Blast and Bacterial blight.

Table 1 presents data regarding the quantity of leaf disease images utilized in this study. The dataset contains four distinct types of leaf diseases. Table 1 displays the distribution of images for each disease category, indicating the number of images accessible for both training and validation purposes.

Number of original images refers to the complete image that is available for each category of leaf disease prior to the data partitioning process. Total for training represents the quantity of images utilized for training models on every category of leaf disease. Total for validation refers to the quantity of images employed to evaluate the performance of models across various types of Leaf Disease.

As an illustration, Tungro disease (class 0) is associated with a total of 1308 unaltered photos. Within the total, 1046 images are designated for training purposes, while 262 images are allocated for validation. The dataset contains a total of 5932 images, out of which 4745 images are used for training and 1187 images are used for validation. Table 1 presents a comprehensive summary of the data distribution used in this investigation pertaining to the detection of leaf diseases.

Table 1 Rice leaf diseases images number

Leaf diseases	Number of original images	Total for training	Total for validation
Tungro (class 0)	1308	1046	262
Brown spot (class 1)	1600	1280	320
Blast (class 2)	1440	1152	288
Bacterial blight (class 3)	1584	1267	317
Total images	5932	4745	1187

2.2 Data Processing

In preprocessing the dataset, this study performed rescale = 1/255 data normalization to ensure that the pixel intensity of the image is converted into a range of 0 to 1. Next, the dataset is divided into training and validation subsets. A total of 20% of the data will be taken as validation data. For training data and validation data, the image size is set to 224 × 224. After going through the division process between training and validation data, the amount of training data is 4747 images belonging to 4 classes and validation data of 1185 images belonging to 4 classes.

2.3 Transfer Learning Base Model

Transfer learning is an approach in the field of artificial intelligence where a model that has been trained for a task can be used as a basis for solving a different related task. In transfer learning, knowledge gained from an initial task (source) can be transferred to a new task (target). The core idea is that understanding gained from one task can help accelerate learning in another task [14].

Image classification can be used in various fields, including chicken farming [15], the difficulty in this research is to get the best accuracy while reducing overfitting in the model to get classification results with the expected accuracy, by using transfer learning the problem in this research can be overcome. Wu et al. proposed a method for transferring gesture data between domains, eliminating the requirement for huge volumes of data to be collected and annotated for each new domain [16].

The base model utilized in this study comprises of sixteen pre-trained models, which are further augmented with the inclusion of the AlexNet model, ultimately resulting in a total of seventeen pre-trained models. All of these pre-trained models are initialized with weights derived from the 'imagenet' dataset. However, the standard fully connected classifier layer that is typically found at the top of such pre-trained models has been omitted. Instead, a customized classifier layer has been added based on the specific requirements of the study. Additionally, the model's architecture has been configured to accommodate image dimensions of 224 × 224 pixels.

Table 2. describes the model layer architecture used in this study. which can be described as follows, the Input Layer, the initial layer that receives input data, images, with a specified dimension of 224 × 224 pixels. Transfer Learning Layer, refers to a set of previously trained models (including an additional AlexNet) with an 'imagenet' weight, which serves as a feature extractor for subsequent layers. Flatten Layer After a transfer learning layer, this layer re-transforms the extracted features into one-dimensional vectors. The dense layer is a fully connected layer that performs classification based on vector features that modify the shape. The last layer of the model is output layer, using the softmax activation function, generates a probability class for the input image.

Table 2 The Models architecture layer

Layer type	Description
Input layer	Initial layer that accepts input data, specifically images, with dimensions specified as 224 × 224 pixels
Transfer learning layer	Refers to the collective set of pre-trained models (including the added AlexNet) with 'imagenet' weights, which serve as a feature extractor for the subsequent layers
Flatten layer	After the transfer learning layers, this layer reshapes the extracted features into a one-dimensional vector
Dense layer	A fully connected layer that performs classification based on the reshaped feature vector
Output layer	The final layer of the model, utilizing softmax activation function, produces class probabilities for the input image

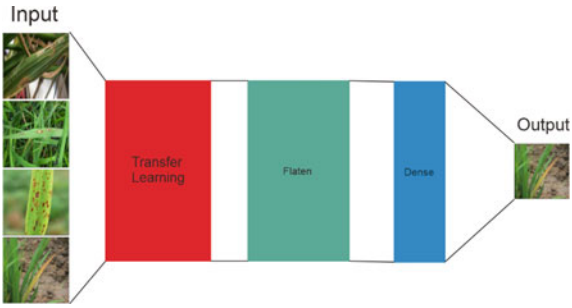


Fig. 3 Base model architecture

The entire base model used as a transfer learning model in this study uses the architectural form shown in Fig. 3 and a complete description of the model name can be seen in Table 3.

Researchers set the scene for training image classification models using multiple transfer learning architectures, using data generators to load images, and measuring accuracy as well as precision and recall metrics during training, this setting also ap-plies to ensemble learning modelling. In detail, all configurations in this study can be seen in Table 4.

2.4 Ensemble Model

Various separate models are combined in an ensemble learning technique to improve generalization ability. Through combining the advantages of deep learning and ensemble learning, ensemble learning models improve generalization ability [17]. The ensemble models include boosting, bagging, stacking, homogeneous or heterogeneous variants, along with explicit or implicit structures, negative correlation-driven deep ensemble models, and decision fusion-orientated deep ensemble models

Table 3 Transfer learning model in this research

No	Model
1	Xception
2	VGG16
3	VGG19
4	ResNet50
5	ResNet101
6	ResNet152
7	InceptionV3
8	MobileNetV2
9	DenseNet201
10	NASNetMobile
11	NASNetLarge
12	EfficientNetB0
13	EfficientNetB5
14	EfficientNetB6
15	EfficientNetB7
16	ConvNeXBase
17	AlexNet

Table 4 Training configuration

Configuration	Value
SEED	42
BATCH_SIZE	64
EPOCHS	50
INPUT_SHAPE	(224,224,3)
Data_Rescaling	1./255
Validation split	0.2
Optimizer	Adam
Loss function	Categorical_crossentropy
Metrics	Accuracy, Precision, Recall
Callbacks	Timing_accuracy_eta_callback

[18], ensemble models can provide high accuracy results on the classification of an image with a small dataset size [19, 20].

The researcher used three transfer learning base models in this study are EfficientNetB0, AlexNet, and MobileNetV2. Because there is one model with poor accuracy results mixed with two models with high accuracy, the resulting ensemble model can produce high accuracy results when these three models are used as basis models in the ensemble model, Fig. 4. shows the ensemble model used in this study.

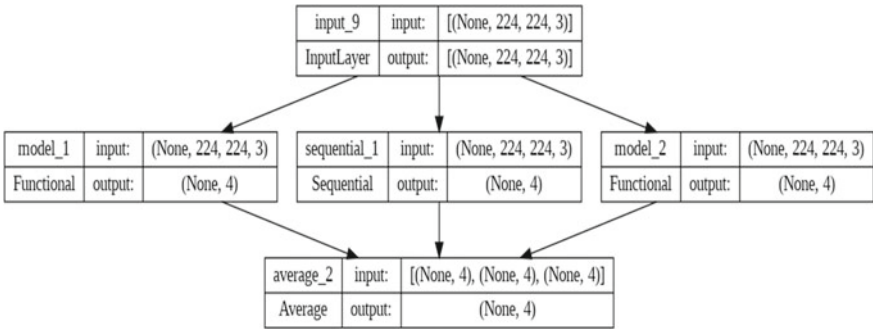


Fig. 4 Ensemble model

Based on the shape of the model in Fig. 4, we can calculate the output results with the equation [21]:

$$output = \frac{1}{N}\sum_{i=1}^{N} Xi \tag{1}$$

where $N =$ is the number of elements in the ensemble, $Xi =$ is the value of the i th element in the ensemble.

EfficientNetB0

The pre-trained EfficientNet was performed by first loading the ImageNet weights into the EfficientNet model followed by the addition of multiple upper layers and fully connected layers for the classification of rice leaf disease types. Several experiments were conducted to analyze the robustness of the fine-tuned EfficientNet compared to the pre-trained models [22].

EfficientNetB0 accepts an image with dimensions of $224 \times 224 \times 3$. The model employs multiple convolutional layers (Conv), a mobile inverse bottleneck Conv (MBConv), and a 3×3 receptive array to extract features from each layer. Compared to other models, the EfficientNetB0 model may scale each dimension using a fixed set of balanced depth, width, resolution, and scaling factors, resulting in a scalable, accurate, and simple-to-build model [23]. Main Features of EfficientNetB0 can be seen in Table 5.

MobileNetV2

MobileNetV2 is one of the convolutional neural network architectures specifically designed for computational efficiency in image processing and object recognition tasks. This architecture was developed by scientists at Google [25] and is an extension of the previous MobileNetV1 model. This model has been used for the performance of ImageNet classification, COCO object detection and image segmentation. The main features of the MobileNetV2 model can be seen in Table 6.

Table 5 Main features of efficientNetB0 [24]

Main features of EfficientNetB0	Description
Compound scaling	EfficientNetB0 leverages "compound scaling," responding proportionally to resolution, depth, and network width. This enables the model to become more powerful and efficient with size increase
Depthwise separable convolu-tion and swish activation	EfficientNetB0 employs the concept of "depth-wise separable convolution" for computational efficiency. Additionally, the model replaces ReLU with "swish" activation, offering smoother non-linearity
MBConv blocks	The EfficientNetB0 architecture utilizes "MBConv" (Mobile Inverted Residual Bottleneck) blocks, combining inverted residuals concept with width-based expansion. This provides a balance between capacity and efficiency
Efficient scaling	This model adopts the "Efficient Scaling" ap-proach that adjusts size-related parameters to enhance performance without compromising efficiency
AutoML approach	EfficientNetB0 is generated through the AutoML (Automated Machine Learning) approach, seeking the best model among various configurations with optimization for accuracy and resources

AlexNet

AlexNet is a groundbreaking CNN architecture that contributed to the resurgence of interest in neural networks and catalyzed the development of deeper and more complex architectures. Its innovations in terms of architecture design, activation functions, normalization techniques, and training strategies laid the foundation for subsequent advancements in the field of deep learning [26].

Rizhevsky introduced the concept of AlexNet, leading to a remarkable leap forward in the realm of computer vision knowledge. The architecture of AlexNet encompasses a total of five convolutional layers, followed by two fully connected layers, aimed at learning distinctive features. Max-pooling is applied after the first, second, and fifth convolutional layers. The network comprises a substantial 630 million connections, 60 million parameters, and a collective 650,000 neurons. AlexNet holds the distinction of being the pioneering example that showcased the potential of employing deep learning in computer vision tasks [27]. Table 7 describes the main features of the AlexNet model.

Table 6 Main feature of MobileNetV2 [25]

Main features of MobileNetV2	Description
Depthwise separable convolution	MobileNetV2 utilizes grouped convolution known as "depthwise separable convolution." This reduces computation through a 1 × 1 convolution step followed by depthwise convolution
Inverted residuals	The MobileNetV2 architecture introduces the "inverted residual" block, which combines dimension expansion and dimension reduction for a balance between expressiveness and efficiency
Bottleneck design	MobileNetV2 implements a "bottleneck" design that reduces the middle layer's dimensions to conserve resources, akin to the concept in the ResNet architecture
Width multiplier and resolution multiplier	MobileNetV2 enables adjustment of the "width multiplier" (number of features per layer) and "resolution multiplier" (image resolution) to control model complexity
Efficiency	MobileNetV2 is designed for use on resource-constrained devices, such as mobile devices. It offers a balance between accuracy and speed

Table 7 Main features of AlexNet [26]

Main features of AlexNet	Description
Early deep learning architecture	AlexNet is one of the widely recognized pioneer convolutional neural network architectures. It was introduced in 2012 by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton
Convolutional and fully connected layers	AlexNet comprises convolutional layers followed by fully connected layers. The architecture consists of 8 layers, including 5 convolutional and 3 fully connected layers
ReLU activation and local response Normali-zation	AlexNet popularized the use of Rectified Linear Unit (ReLU) activation as a non-linear activation function in its layers. It also introduced the technique of Local Re-sponse Normalization
overlapping pooling	AlexNet utilizes max pooling technique within its convo-lutional layers, with some pooling layers overlapping to enhance feature enrichment
Dropout and data Aug-mentation	AlexNet introduced the utilization of dropout to miti-gate overfitting during training. Additionally, it imple-ments data augmentation to expand the training da-taset

2.5 System Evaluation

To provide a comprehensive overview of the model's performance in each class, while considering the distribution of instances in each class and the overall performance of the model, the researcher presents the Classification Report in the following detailed form:

Precision: Precision measures the proportion of positives correctly identified out of the total positives identified by the model.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{2}$$

Recall: Recall (Sensitivity) measures the proportion of correctly identified positives out of the total true positives.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{3}$$

F1-Score: F1-score is the harmonic mean of precision and recall. It provides a balance between precision and recall.

$$F1 - Score = 2x\frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

Support: Support is the total number of instances in a class. It helps provide context to the other metrics.

Accuracy: Accuracy measures how correctly the model classifies all instances.

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Instances} \tag{5}$$

Macro Avg: Macro average calculates the average of the evaluation metrics for each class.

$$MacroAvg = \frac{M_1 + M_2 + \cdots + M_n}{N} \tag{6}$$

Weighted Avg: Weighted average is a weighted average of the evaluation metrics based on the class distribution.

$$\text{WeightedAvg} = \frac{S_1 x M_1 + S_2 x M_2 + \cdots + S_n x M_n}{TotalInstance} \tag{7}$$

where: N is the number of classes, M is the evaluation metric, S is the number of instances in the class.

Meanwhile, this research uses Python 3 Google Compute Engine Backend (GPU) with specifications of 12.7 GB RAM, 15.0 GB GPU and 166.GB Disk, Hardware Accelerator T4 GPU with Shape High RAM.

3 Result

Classification of rice leaf diseases in this study was carried out by taking a comparison with previous research [6] with the same research environment atmosphere such as the number of epochs of the dataset used so that the results of the research were obtained, Table 8 show the results of the research.

Table 8 Comparison of results with previous research

No	Model	Previous research	Training time	This research	Training time (seconds)
1	Xception	0.79	392.726223630	0.96	557.06
2	VGG16	0.79	385.728259577	0.96	602.89
3	VGG19	0.78	388.040710733	0.98	463.06
4	ResNet50	0.78	385.395347467	0.90	1572.23
5	ResNet101	0.79	391.797446297	0.91	1734.28
6	ResNet152	–	–	0.86	3524.43
7	InceptionV3	0.79	385.660891293	0.97	78.70
8	MobileNetV2	–	–	0.99	223.17
9	DenseNet201	–	–	0.99	376.57
10	NASNetMobile	–	–	0.98	192.13
11	NASNetLarge	–	–	0.95	412.49
12	EfficientNetB0	–	–	0.38	783.01
13	EfficientNetB5	–	–	0.31	3552.56
14	EfficientNetB6	–	–	0.46	3613.60
15	EfficientNetB7	–	–	0.27	8017.06
16	ConvNeXBase	–	–	0.92	258.28
17	AlexNet	0.79	381.799865173	0.91	767.97

3.1 Transfer Learning Model

Classification of rice leaf diseases in this study was carried out by taking a comparison with previous research with the same research environment atmosphere such as the number of epochs of the dataset used so that the results of the research were obtained.

This study added several transfer learning models that were not tested in previous studies. The results showed in Table 8 that not all models have high accuracy when used for training the same rice leaf disease classification data. Models that have low accuracy values in this study are used as base models in ensemble learning.

3.2 Ensemble Learning Model

Researchers conducted several combinations of base models, one base model was taken from the model with the lowest accuracy and a small architecture, choosing this small architecture to consider the limited memory of the GPU during training data. Then two models with the highest value were also selected to cover the weaknesses of the base model with low accuracy.

Model Stage

In the formation of this ensemble model, researchers made two models. The first model is a combination of EfficientNetB0 and AlexNet. The second model is a combination of EfficientNetB0, AlexNet, and MobileNetV2.

Table 9. shows that in the first stage, we built an ensemble model by combining EfficientNetB0 and AlexNet architectures. The training results show an accuracy of 0.35 with a training time of 783.25 s. In the second stage, researchers continued by building a more complex ensemble model by combining EfficientNetB0, AlexNet, and MobileNetV2 architectures. This ensemble model achieves an accuracy of 0.97 with a training time of 36.42 s.

From these results, the second model, which involves a combination of three architectures, provides a significant improvement in accuracy with a much shorter training time than the first model that only combines two architectures. This shows that combining multiple architectures in an ensemble model can result in better performance in image classification tasks.

Table 9 Ensemble model stage

Model stage	Used architectures	Accuracy	Training time (seconds)
First	EfficientNetB0 + AlexNet	0.35	783.25
Second	EfficientNetB0 + AlexNet + MobileNetV2	0.97	36.42

Classification Report

This section presents the classification report for the developed classification model. The classification report is a comprehensive evaluation tool, that provides insights into the model's performance across various classes. It includes precision, recall, F1-score, and Support metrics, which offer a detailed understanding of the model's ability to classify instances belonging to different classes correctly.

Table 10 presents the classification report as an evaluation of the second model's performance. This model achieved an overall accuracy of 0.97, indicating that it correctly classified 97% of the instances in the dataset. The report provides insights into each class's precision, recall, and F1-score.

For class 0, the precision is 1.00, which means that when the model predicts an instance as class 0, it is correct 100% of the time. The recall is 0.97, indicating that the model identified 97% of the actual instances belonging to class 0. The F1-score, a harmonic mean of precision and recall, is 0.98, representing this class's balance between precision and recall.

For class 1, the model achieved a precision of 0.97, correctly classifying 97% of the instances it predicted as class 1. The recall is 0.98, signifying that the model captured 98% of the actual instances of class 1. The F1-score for this class is 0.97, reflecting a strong balance between precision and recall.

Class 2 exhibits a precision of 0.91, indicating that the model accurately predicted 91% of the instances as class 2. The recall for class 2 is 1.00, implying that the model correctly identified all instances of class 2. The F1-score stands at 0.95, indicating a good balance between precision and recall.

Class 3's precision is 1.00, signifying that the model correctly identified all instances it predicted as class 3. The recall, at 0.91, suggests that the model captured 91% of the true instances of class 3. The F1-score for this class is 0.95, indicating a harmonious balance between precision and recall.

The support column indicates the count of instances in each class. For instance, class 0 has a support of 316, class 1 has 288 instances, class 2 has 320 instances, and class 3 has 261 instances. This information helps contextualize the performance metrics by considering the distribution of instances across classes.

Table 10 Classification of second stage model

Class	Precision	Recall	F1-Score	Support
0	1.00	0.97	0.98	316
1	0.97	0.98	0.97	288
2	0.91	1.00	0.95	320
3	1.00	0.91	0.95	261
Accuracy	–	–	0.97	1185
Macro Avg	0.97	0.96	0.97	1185
Weighted Avg	0.97	0.97	0.97	1185

The macro average F1-score, which calculates the average F1-score across all classes, is 0.97, reflecting the overall model performance across classes. The weighted average F1-score, which accounts for class imbalance, is also 0.97.

4 Discussion

This research comprehensively evaluates various deep learning architectures for image classification tasks. Tables displaying previous research performance, training time, and results of this study underline the significant progress achieved. Comparisons across several models show a pattern where advanced architectures tend to achieve higher accuracy. The models showed excellent accuracy, although some showed low accuracy below 0.40, with highly variable training times.

In the ensemble learning development that is shown in Table 9., the two-stage model, combining EfficientNetB0 with AlexNet and later with MobileNetV2, exemplifies an approach to utilize the strengths of different architectures. This strategy significantly improved accuracy, as evidenced by the marked improvement between the first and second stages. This shows the importance of model composition and signifies that combining different models can lead to better classification results.

Although the proposed research shows significant progress in image classification using various deep learning architectures, it is important to recognize certain limitations inherent to such approaches. Firstly, selecting architectures and their incorporation may not be universally optimal for all datasets and applications. The effectiveness of ensemble methods, although proven in this study, can still show variations based on dataset characteristics and domain-specific requirements. In addition, the rapid evolution of deep learning architectures suggests that newer models may emerge that could outperform the models evaluated in this study, thus requiring continuous updating and adaptation.

Another limitation is related to computational resources. Some high-performance architectures, such as EfficientNetB6 and EfficientNetB7, require significant computing power and time for training. This may hinder practicality in resource-constrained scenarios. The training time for some architectures is much longer, which may hinder their applicability in real-time applications where low-latency response is critical.

The implications of the proposed research hold significant promise for both the field of deep learning and practical applications. Firstly, the successful fusion of multiple architectures in the ensemble model underscores the potential for harnessing the strengths of different models to create more accurate and versatile classifiers. This finding encourages further exploration of ensemble techniques in various machine learning tasks beyond image classification. Researchers and practitioners could adopt a similar approach to tackle complex problems, potentially leading to improved performance and generalization across domains.

Moreover, the substantial accuracy improvements achieved through the ensemble approach could have far-reaching implications in real-world applications. For

instance, industries relying on image-based quality control, such as manufacturing and agriculture, could benefit from more accurate defect detection and product evaluation. Additionally, the enhanced accuracy of medical image classification achieved in this research could translate to more reliable disease diagnosis and early intervention, potentially revolutionizing the healthcare sector. The demonstrated success of combining deep learning architectures could inspire collaboration between experts from diverse fields to develop tailored models that address specific challenges effectively.

5 Conclusion

This research has demonstrated the profound impact of employing ensemble methods that combine diverse deep-learning architectures in the context of image classification. The results have been nothing short of remarkable, showcasing a remarkable increase in accuracy from a baseline of 0.35 to an impressive 0.97 in the ensemble model. This achievement underscores the immense potential of leveraging different architectural strengths and their complementary features to enhance classification accuracy substantially. The findings affirm the critical importance of thoughtful model selection and ensemble strategies in pursuing superior classification performance. Furthermore, this research is a testament to computer vision's ongoing evolution and innovation. As we look ahead, future research avenues may delve into the fine-tuning of hyperparameters, expanding the study's scope to encompass other challenging computer vision tasks and exploring strategies for bias mitigation and model interpretability. Ultimately, this study provides valuable insights and practical solutions that hold promise for researchers and practitioners seeking to navigate the complex landscape of image classification challenges.

References

1. Val-Torregrosa B, Bundó M, San Segundo B (2021) Crosstalk between nutrient signalling pathways and immune responses in rice.https://doi.org/10.3390/agriculture11080747
2. Ramesh S, Vydeki D (2020) Recognition and classification of paddy leaf diseases using optimized deep neural network with Jaya algorithm. Inf Process Agric 7:249–260. https://doi.org/10.1016/j.inpa.2019.09.002
3. Dogra R, Rani S, Singh A, Albahar MA, Barrera AE, Alkhayyat A (2023) Deep learning model for detection of brown spot rice leaf disease with smart agriculture. Comput Electr Eng, 109. https://doi.org/10.1016/j.compeleceng.2023.108659
4. Ahad MT, Li Y, Song B, Bhuiyan T (2023) Comparison of CNN-based deep learning architectures for rice diseases classification. Artif Intell Agric 9:22–35. https://doi.org/10.1016/j.aiia.2023.07.001
5. Sarkar C, Gupta D, Gupta U, Hazarika BB (2023) Leaf disease detection using machine learning and deep learning: review and challenges. Appl Soft Comput, 110534. https://doi.org/10.1016/j.asoc.2023.110534

6. Sethy PK, Barpanda NK, Rath AK, Behera SK (2020) Deep feature based rice leaf disease identification using support vector machine. Comput Electron Agric, 175. https://doi.org/10.1016/j.compag.2020.105527
7. Julianto A, Sunyoto A (2021) A performance evaluation of convolutional neural network architecture for classification of rice leaf disease. IAES Int J Artif Intell 10:1069–1078. https://doi.org/10.11591/IJAI.V10.I4.PP1069-1078
8. N K, Narasimha Prasad LV, Pavan Kumar CS, Subedi B, Abraha HB, Sathishkumar VE (2021) Rice leaf diseases prediction using deep neural networks with transfer learning. Environ Res., 198. https://doi.org/10.1016/j.envres.2021.111275
9. Shrivastava VK, Pradhan MK, Minz S, Thakur MP (2019) Rice plant disease classification using transfer learning of deep convolution neural network. In: International archives of the photogrammetry, remote sensing and spatial information sciences—ISPRS archives, pp 631–635. International society for photogrammetry and remote sensing. https://doi.org/10.5194/isprs-archives-XLII-3-W6-631-2019
10. Bhattacharya S, Mukherjee A, Phadikar S (2020) A deep learning approach for the classification of rice leaf diseases. In: Advances in Intelligent systems and computing, pp 61–69. Springer. https://doi.org/10.1007/978-981-15-2021-1_8
11. Bhowmik A, Sannigrahi M, Chowdhury D, Das D (2022) RiceCloud: a cloud integrated ensemble learning based rice leaf diseases prediction system. In: INDICON 2022–2022 IEEE 19th India council international conference. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/INDICON56171.2022.10039790
12. Sembiring A, Away Y, Arnia F, Muharar R (2023) The performance of various concise convolutional neural network configurations in classifying tomato diseases based on leaf images. In: Triwiyanto T, Rizal A, Caesarendra W (eds) Proceeding of the 3rd international conference on electronics, biomedical engineering, and health informatics, pp 373–389. Springer Nature Singapore, Singapore
13. Triwiyanto T, Lusiana, Wakidi LF, Amrinsani F (2022) Deep learning approach to detect the Covid-19 Infection using Chest X-ray image: a review. In: Triwiyanto T, Rizal A, Caesarendra W (eds) Proceedings of the 2nd international conference on electronics, biomedical engineering, and health informatics, pp 237–251. Springer Nature Singapore, Singapore
14. Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, Xiong H, He Q (2021) A comprehensive survey on transfer learning.https://doi.org/10.1109/JPROC.2020.3004555
15. Junaidi A, Adhinata FD, Iskandar AR, Lasama J (2022) Image classification for egg incubator using transfer learning VGG16 and InceptionV3. In: Triwiyanto T, Rizal A, Caesarendra W (eds) Proceedings of the 2nd international conference on electronics, biomedical engineering, and health informatics, pp 85–95. Springer Nature Singapore, Singapore
16. Wu BX, Yang CG, Zhong JP (2021) Research on transfer learning of vision-based gesture recognition. Int J Autom Comput 18:422–431. https://doi.org/10.1007/s11633-020-1273-9
17. Zeynali M, Seyedarabi H, Afrouzian R (2023) Classification of EEG signals using transformer based deep learning and ensemble models. Biomed Signal Process Control, 86. https://doi.org/10.1016/j.bspc.2023.105130
18. Mohammadi A, Shaverizade A (2021) Ensemble deep learning for aspect-based sentiment analysis. Int J Nonlinear Anal Appl 12:29–38. https://doi.org/10.22075/IJNAA.2021.4769
19. Phung VH, Rhee EJ (2019) A High-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets. Appl Sci (Switzerland) 9. https://doi.org/10.3390/app9214500
20. Cetinoglu YK, Koska IO, Uluc ME, Gelal MF (2021) Detection and vascular territorial classification of stroke on diffusion-weighted MRI by deep learning. Eur J Radiol, 145. https://doi.org/10.1016/j.ejrad.2021.110050
21. Kunapuli G (2023) Ensemble methods for machine learning
22. Zulfiqar F, Ijaz Bajwa U, Mehmood Y (2023) Multi-class classification of brain tumor types from MR images using EfficientNets. Biomed Signal Process Control, 84. https://doi.org/10.1016/j.bspc.2023.104777

23. Ali MS, Hassan A, Rahim A, Ashraf MH, Rahim A, Saghir S (2023) Motor imagery EEG classification using fine-tuned deep convolutional EfficientNetB0 model. In: 3rd IEEE international conference on artificial intelligence, ICAI 2023, pp 1–6. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICAI58407.2023.10136681
24. Tan M, Le QV (2019) EfficientNet: rethinking model scaling for convolutional neural networks
25. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C (2018) MobileNetV2: inverted residuals and linear bottlenecks
26. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks
27. Taye MM (2023) Theoretical understanding of convolutional neural network: concepts, architectures, applications, future directions. https://doi.org/10.3390/computation11030052

Improving Unbalanced Security X-Ray Image Classification Using VGG16 and AlexNet with Z-Score Normalization and Augmentation



Diao Qi, Apri Junaidi, Chan Weng Howe, and Azlan Mohd Zain

Abstract Addressing the challenge of unbalanced data sets in convolutional neural network (CNN) models for image recognition, this study aims to investigate the impact of data augmentation and normalization. The problem lies in the limited generalization of the model due to class-based data differences, hence the need for several techniques in data preprocessing such as data augmentation and normalization. The main contribution of this research is a comprehensive analysis of the effectiveness of data augmentation and normalization techniques in improving model performance. The research utilized the AlexNet and VGG16 architectures and conducted extensive experiments on data sets with varying degrees of imbalance. Data augmentation generates additional examples, while normalization reduces convergence issues. The results show that training the AlexNet model without augmentation results in a low accuracy of 0.24, underscoring the challenges posed by skewed data distributions. In contrast, augmented data substantially improves performance, with AlexNet achieving an accuracy of 0.91 and VGG16 achieving 0.84. In addition, normalized data also made a positive contribution, showing an accuracy of 0.74 for AlexNet and 0.67 for VGG16. In conclusion, data augmentation and normalization techniques are essential in reducing the effects of data imbalance, thereby improving the generalizability of the models. The improved accuracy of the data using normalization techniques indicates the ability of the model to read the data after normalization. This study underscores the importance of preprocessing strategies in optimizing model performance and advancing the field of deep learning in image recognition.

Keywords Unbalanced dataset · Image classification · X-Ray security image · Normalisation

D. Qi · A. Junaidi (✉) · C. W. Howe · A. M. Zain
Faculty of Computing, Universiti Teknologi Malaysia, UTM Johor Bahru, Johor 81310, Malaysia
e-mail: junaidi20@graduate.utm.my

D. Qi
Faculty of Artificial Intelligence, Zhejiang Dongfang Polytechnic, Wenzhou 32500, China

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_14

1 Introduction

Imbalanced data occurs when there is a significant disparity in the distribution of classes or labels within a dataset. In this context, one group is disproportionately represented compared to the others, leading machine learning models to perform poorly in recognizing minority classes [1]. This condition is often encountered in various applications such as financial fraud detection, rare medical diagnoses, and anomaly detection, where the less common class can have crucial implications despite its infrequent occurrence in the dataset [2]. Solutions to address imbalanced data include utilizing oversampling, under sampling, data augmentation or class-based approaches that focus on enhancing the recognition of minority classes during the learning process [3].

Image classification for unbalanced datasets can be challenging due to a large amount of information irrelevant to the outcome of the classifier's training and the imbalance in the number of samples across different classes [4]. Various techniques have been proposed to address this issue, such as rebalancing the datasets through oversampling and/or under sampling using machine learning classifiers [5]. Additionally, deep learning models have been used to improve image classification performance on unbalanced datasets. For example, a depth convolution neural network has been developed to classify category imbalance images, achieving better classification results for small-class samples in unbalanced datasets [6]. Another approach involves using a bias prediction network to recover biases from extracted features, resulting in improved performance in few-shot image classification on biased datasets [7, 8]. These methods provide practical guidance for training viable classifiers in the context of image classification for unbalanced datasets.

Figure 1 illustrates a graph of data imbalance. In this dataset, there are two data classes that have different amounts of data. This imbalance refers to a situation where one class has significantly more samples than the other. In this case, the class that has a smaller number of samples is considered the minority class, while the class with a larger number of samples is considered the majority class.

This kind of data imbalance can have a negative impact on the performance of the classification model. When a model is trained on an imbalanced dataset, it tends to favour the majority class because there are more samples representing that class. As a result, the minority class often does poorly in the predictions. For example, if we have an image dataset with 90% Nondangerous images and only 10% dangerous images, the model tends to be better at recognizing dangerous images than dangerous images [9]. This can lead to biased accuracy and may not accurately reflect the success of classifying the minority class. Therefore, it is crucial to understand and address data imbalance in the classification process, such as using oversampling or under sampling. techniques, to enhance the model's performance in classifying the minority class [10].

Alaba et al. [11] present a deep learning model for fish species recognition using object detection that combines MobileNetv3-large and VGG16 backbone networks with an SSD detection head. The problem of unbalanced data sets in fish species

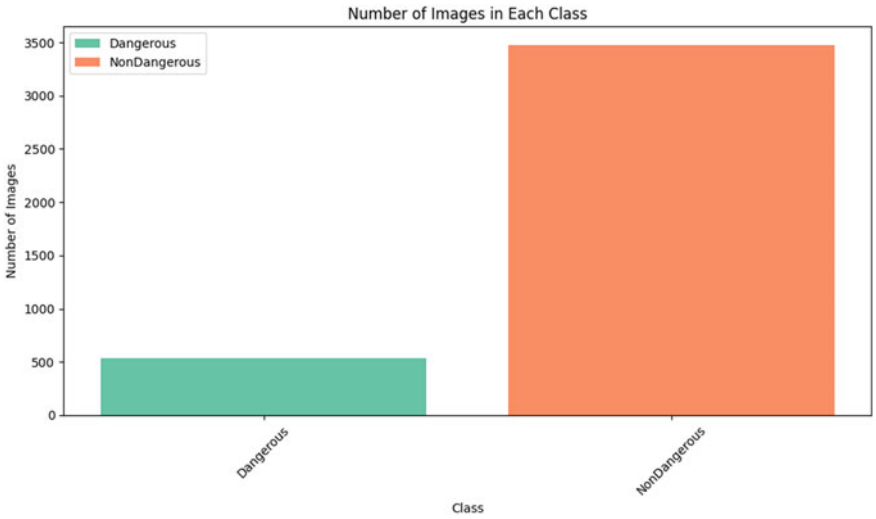


Fig. 1 Dataset distribution of Imbalance data

recognition is addressed in this work. To address this imbalance, this work presents a class-aware loss function that takes the number of cases in each species into consideration and provides more weight to species with fewer examples. According to the experimental data, class-aware loss improves model performance by 79.7%. This paper does not present the results of comparisons between previous research and the results obtained, making it challenging to measure achievements based on previous research. This study [12] addresses the problems of early potato disease classification and data imbalance by modifying the GhostNet model and applying techniques such as ReLU activation, Dropout, and LS-CCE loss function. The experimental results show a significant improvement in potato disease recognition accuracy. The results show that the recognition accuracy reaches 93.27%. Nonetheless, further analyses may be required to ensure the reliability and generality of this approach in various situations. Sun et al. [13] successfully addresses the significant problems in melanoma classification: data imbalance and low accuracy. By using data expansion, sampling, and ensemble learning methods, the authors significantly improved the classification accuracy from 79.9% to 97.8%. This finding can potentially contribute to skin disease detection and classification using computer vision approaches. However, as always, more trials and analyses may be needed to understand the strengths and limitations of this approach in more depth. This research [14] creates a novel approach for automatic threat recognition and detection in X-ray scans of luggage. The proposed system combines CNN and BLS with automatic design adaptation that enables recognizing hazardous objects in imbalanced data. Experimental results show that the system outperforms previous methods regarding detection accuracy by 0.9489. These findings can bring significant developments in security and threat detection technologies in various fields that require image

understanding, such as airports, shopping center, and shipping. While this research provides improved performance in threat detection, there are questions about the extent to which these models will perform well in natural environments. Factors such as lighting variations, object position, and shape variations may affect the model's ability to detect threats accurately in real scenarios. Xu et al. [15] introduces a novel approach for fish species identification by combining SE-ResNet152 architecture and class-balanced focal loss. This illustrates the novelty in fish species identification efforts using deep learning techniques. The test results show that the proposed method has a high accuracy of 98.80%, 96.67%, and 91.25% in identifying fish species through various image views, even outperforming other standard methods. This research overcomes the imbalance problem in the number of samples in different fish species categories through class-balanced focal loss. This helps in improving the classification performance for underrepresented categories. The proposed method is tested on the Fish-Pak dataset, which may have specific characteristics irrelevant to other datasets. The performance of this method on other datasets needs to be further evaluated to understand the extent of its generalizability. Sambasivam and Opiyo [16] develops a Convolutional Neural Networks (CNN) model to overcome the problem of class imbalance and poor performance on cassava disease datasets. The techniques used to overcome these problems include the use of class weights, focal loss, and Synthetic Minority Over-sampling Technique (SMOTE) as well as variations in input image dimensions. Using class weights, focal loss, SMOTE, and larger input image dimensions showed positive results. Accuracy increased by more than 5%, and log loss decreased drastically to 0.06% from more than 20% when the class imbalance solving technique was used, along with additional data and larger input image dimensions. The strength of this research is its comprehensive approach to addressing the issues of class imbalance and small dataset size. This research provides a real solution in developing a model that can be used in mobile applications to assist farmers in detecting diseases in cassava plants more efficiently. The drawback of this research is that although this research provides positive results in overcoming the problem of class imbalance, further research is still needed to address this concern, it is important to clarify the primary objects of the X-ray images used in this study. The X-ray images in our research represent two main categories: 'dangerous' and 'nondangerous' images, specifically related to scenarios encountered at airports. These categories are utilized to identify situations where multiple diseases coincide on the same plant, which is a common challenge in agricultural pathology. Furthermore, it is essential to note that our model's applicability to more complex field situations still require further testing and validation. In dealing with the issue of unbalanced data, our research primarily employs data augmentation by increasing the number of images in classes with a limited number of samples. Additionally, we utilize Z-score normalization as a preprocessing step. Through our research, we have demonstrated that these techniques significantly enhance accuracy when dealing with unbalanced data conditions.

2 Materials and Method

The X-ray security image classification model at the airport is depicted in Fig. 2. This figure illustrates the proposed methodology discussed in this study, comprising two distinct approaches. The left chart elucidates the treatment of unbalanced data through Data Augmentation, while the chart on the right delineates the process of handling data using standard data normalization techniques.

Figure 2 depicts a three-step workflow that begins with dataset collection and progresses to the first stage, which is focused on improving data performance through preprocessing. The key difficulty addressed in this research during this phase is data imbalance. Following that, the second stage attempts to achieve maximum accuracy through a series of procedures that include the formation of a basic model configuration, model construction, and model fitting. Finally, model training is performed to attain accuracy with unbalanced datasets. The primary difference between these two models is in the early step, where one undergoes data augmentation using augmentation techniques to increase the amount of the dataset, and the other uses z-score normalization for data processing.

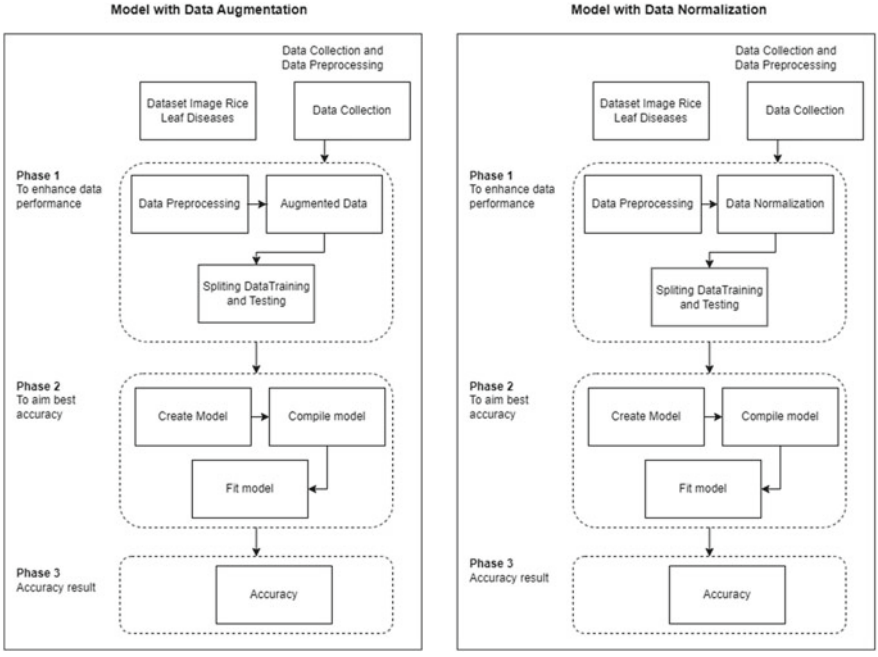


Fig. 2 The proposed method

2.1 Dataset

The dataset used in this study was obtained on the web page https://universe.roboflow.com/airport-security-scanning/xray-upzdb. This dataset is an Open Source Dataset and has two classes Dangerous and NonDangerous [17]. The dataset originally consisted of three folders, namely, Train, Valid, and Test. Still, to facilitate preprocessing, all data in the three folders was combined based on the class so that the amount of data per class was obtained as follows, Dangerous class has 538 images and NonDangerous class has 3475 images. Figures 3 and 4 describe the images of the NonDangerous and Dangerous classes.

There are three phases starting from collecting datasets and then entering the first phase which is to improve data performance by preprocessing data. In this process there is the core problem in this research, namely data unbalance. Then stepping on the second phase to aim for the best accuracy by going through several steps, namely making the base model configuration, compiling the model, and fitting the model. The last stage gets the accuracy of the model that has been trained with the unbalanced dataset.

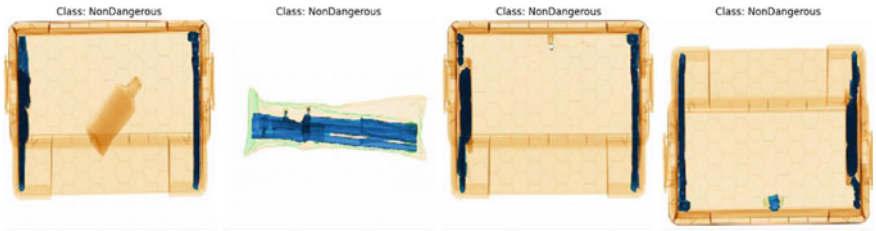


Fig. 3 NonDangerous class



Fig. 4 Dangerous class

2.2 Data Augmentation

Data augmentation is a technique used in data processing and machine learning to increase the amount and variety of training data by manipulating existing data [18]. The main goal of data augmentation is to improve the ability of machine learning models to recognize more common and varied patterns [19]. Zhang discusses the algorithmic implications of data augmentation in the context of robust transfer learning [20].

Zhang proposes an effective method to augment expert demonstrations in behavioral cloning to alleviate the problem of compounding error. Rihal and Assal present an effort to develop learning models based on data augmentation for the digital documentation of heritage structures [21]. Liu et al. propose a general data augmentation technique for neural machine translation to address the challenge of data sparsity [22]. In this study, the amount of data is added to classes that have little data, Fig. 5. Explain about the augmented data result.

In this study, we employed a robust data augmentation technique to enhance the diversity and volume of our dataset. We utilized an ImageDataGenerator with various augmentation parameters, including a rotation range of 20 degrees, width and height shift ranges of 0.2, shear range of 0.2, zoom range of 0.2, and horizontal flipping, along with the 'nearest' fill mode for image augmentation. These augmentation settings were implemented to introduce variations in the training data, such as altering perspectives, orientations, and scales of the images. The objective was to improve the model's ability to generalize and perform effectively on a wider range of real-world scenarios. The results of the implemented data augmentation are presented in the subsequent sections, illustrating the positive impact on the model's performance in handling the challenges posed by unbalanced datasets.

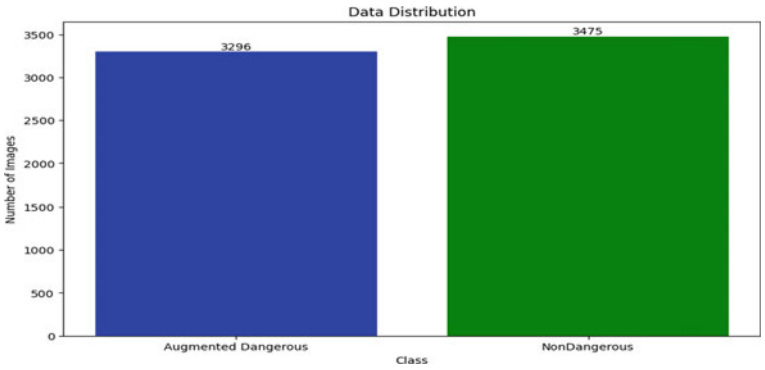


Fig. 5 Data augmentation result

2.3 Data Normalization

Data normalization, also known as data scaling, is the process of rescaling the values in a feature (variable) in a dataset so that the values have a more uniform range or fall within a certain, more defined range. The goal is to ensure that features with different ranges of values have an equal impact on machine learning models or data analysis algorithms [23]. The field of data science, which encompasses the creation of techniques for capturing, storing, and examining data, is intricately linked with data normalization. This is due to its broad scope, which includes various forms of data analysis, incorporating the utilization of machine learning algorithms and dimensionality reduction techniques [24]. There are several standard methods for normalising data, in this study using the Z-Score Normalisation (Standardisation) method: This method changes the data in such a way that it is averaged (μ) to 0 and the standar deviation (σ) to 1, the equation seems in Eq. (1):

$$x_{new} = \frac{x - \mu}{\sigma} \tag{1}$$

where $x_{new} =$ is the normalized value, x is the original value, μ is the average value in the feature, dan σ is the standard deviation of the values in the feature.

2.4 AlexNet

AlexNet is a convolutional neural network (CNN) architecture that has been very influential in the development of deep learning, especially in the field of image recognition. AlexNet was first introduced by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton in 2012 as part of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competition, where they won with impressive results [25].

AlexNet, a pioneering convolutional neural network (CNN), brought about significant advancements in image recognition. It introduced a deeper layer architecture compared to prior models, featuring 8 convolutional layers and incorporating elements such as pooling, contrast normalization, and fully connected layers. By employing convolutional layers as feature extractors and subsequently using max-pooling layers to curb overfitting, AlexNet optimized data analysis. The use of Rectified Linear Activation (ReLU) in nearly all convolutional layers helped overcome gradient vanishing issues and expedited training [26].

Addressing overfitting, AlexNet utilized dropout in fully connected layers, randomly deactivating neuron units during training to enhance tolerance to variations. It also adopted data augmentation techniques, diversifying training data to prevent overfitting [27]. Additionally, AlexNet's architecture originally accommodated parallel processing using two GPUs, with similar architectures on each side managing distinct data portions. The combined outcomes from both sides contributed

to its remarkable success, setting the stage for subsequent deep CNN models like VGG, GoogLeNet, and ResNet [28].

2.5 Vgg16

VGG16 is one of the well-known convolutional neural networks (CNN) architectures in image recognition and image processing. It was developed by a team of researchers at the Visual Geometry Group (VGG) at the University of Oxford. VGG16 is known for the depth and uniformity of its convolutional layer structure [29].

VGG16 is renowned for its ability to extract features on images, but it also has drawbacks in terms of complexity and higher computational requirements. Although there have been more advanced models since then, VGG16 remains an important foundation in the development of more complex CNN architectures [30].

2.6 System Evaluation

In this study the authors used the Accuracy metric, which is one of the evaluation metrics used to measure the extent to which the classification model can correctly predict the class or label of the test data. It gives an idea of how accurate the model is in distinguishing between different classes, the metric is described in the Eq. (2).

$$Accuracy = \frac{Number of Correct Predictions}{Total Number of Data tested} \tag{2}$$

The equation provided represents the calculation of accuracy, a critical metric in our research for evaluating the performance of the classification model. In our context, accuracy measures how well the model correctly predicts outcomes. It is computed by dividing the number of correct predictions by the total number of data points tested. For instance, when assessing our model's ability to identify dangerous items in airport security images, 'Number of Correct Predictions' refers to the images correctly identified as dangerous, while 'Total Number of Data tested' corresponds to the total number of images evaluated.

3 Result and Discussion

Several tests have been conducted in this study. Initially, the data was tested using the AlexNet model without the data augmentation and data normalization processes. In the subsequent stage, testing was performed using data that had been augmented and normalized.

Table 1 Experimental results

Model	Data augmentation	Normalization	Accuracy
AlexNet	No	No	0.24
AlexNet	Yes	No	0.91
AlexNet	No	Yes	0.74
VGG16	No	No	0.24
VGG16	Yes	No	0.84
VGG16	No	Yes	0.67

3.1 Training Without Augmentation and Normalization

Training data without this augmentation, using the AlexNet model produces a very low accuracy of 0.24, until this stage it is necessary to augment the data, the low accuracy is produced because of unbalanced data, and the model only learns on classes that have large data.

3.2 Training with Augmentation and Normalization

Training with augmentation and normalization data, can produce high accuracy, for training with augmented data for AlexNet produces an accuracy of 0.91, VGG16 produces 0.84. For training using Normalized Data AlexNet produces an accuracy of 0.74 while for VGG16 produces an accuracy of 0.67. From the explanations provided in Sects. 3.1 and 3.2, the results of our experiments can be comprehensively presented in Table 1

3.3 Discussion

The interpretation of the study's findings reveals the critical role played by data augmentation and normalization in enhancing the performance of convolutional neural network (CNN) models. The initial accuracy of merely 0.24, observed when training the AlexNet model without augmentation, starkly illustrates the challenges posed by imbalanced datasets. This underscores the necessity of addressing data imbalance to enable models to effectively learn from underrepresented classes.

A significant boost in accuracy is demonstrated through the implementation of augmentation and normalization techniques. Training with augmented data results in substantial accuracy gains, reaching 0.91 for AlexNet and 0.84 for VGG16 models. Similarly, normalized data contributes positively, with AlexNet achieving an accuracy of 0.74 and VGG16 an accuracy of 0.67. These comparative results highlight the effectiveness of both techniques in improving model generalization and class-wise

Table 2 Accuracy comparison with previous studies

Model	Accuracy in thia study	Accuracy in [31]	Accuracy in [32]
AlexNet	0.91	–	0.63
VGG16	0.84	0.68	–

performance. When comparing the accuracy achieved with data augmentation in this study to the previous research, the results can be seen in Table 2.

While the study sheds light on the advantages of data augmentation and normalization, certain limitations should be acknowledged. The augmentation process might introduce noise or unrealistic instances, impacting the model's performance in real-world scenarios. Moreover, the impact of varying augmentation strategies on model robustness requires further exploration. Additionally, while normalization aids convergence, it may not always be suitable for specific data distributions, warranting careful consideration.

The implications of the findings are substantial. Incorporating data augmentation and normalization is pivotal, particularly for models trained on imbalanced datasets. Augmentation not only rectifies the data distribution but also bolsters the model's ability to generalize across diverse classes. The study underscores the significance of preprocessing techniques in deep learning pipelines, highlighting their potential to bridge the gap between theoretical model architectures and real-world performance.

4 Conclusion

In conclusion, this study has shed light on the significant impact of data augmentation and normalization techniques on the performance of convolutional neural network (CNN) models. Our findings have highlighted the inherent challenges posed by imbalanced datasets, as exemplified by the low accuracy of 0.24 achieved when training the AlexNet model without augmentation. This underscores the imperative nature of employing data augmentation techniques to rectify such imbalances and enhance the model's ability to generalize across diverse classes.

The experimental results have provided strong support for the effectiveness of data augmentation and normalization. The substantial improvements in accuracy, with AlexNet and VGG16 models achieving 0.91 and 0.84 accuracy, respectively, when trained with augmented data, and 0.74 and 0.67 accuracy when trained with normalized data, underscore the potential of these techniques to enhance model performance across various scenarios.

The implications of these findings are significant. The incorporation of data augmentation and normalization is pivotal, particularly for models trained on imbalanced datasets. Augmentation not only rectifies the data distribution but also

strengthens the model's ability to generalize across diverse classes. This study emphasizes the importance of preprocessing techniques in deep learning pipelines, highlighting their capacity to bridge the gap between theoretical model architectures and real-world performance.

References

1. Kumar P, Bhatnagar R, Gaur K, Bhatnagar A (2021) Classification of imbalanced data: review of methods and applications. IOP Conf Ser Mater Sci Eng 1099:012077. https://doi.org/10.1088/1757-899x/1099/1/012077
2. Rekha G, Tyagi AK, Sreenath N, Mishra S (2021) Class imbalanced data: open issues and future research directions. In: 2021 International conference on computer communication and informatics, ICCCI 2021. Institute of Electrical and Electronics Engineers Inc. (2021). https://doi.org/10.1109/ICCCI50826.2021.9402272
3. Komori O, Eguchi S (2019) Introduction to imbalanced data. Presented at the (2019).https://doi.org/10.1007/978-4-431-55570-4_1
4. Gusmão G, Raposo A, Oliveira R, de, Barbosa C (2022) Treating dataset imbalance in fetal echocardiography classification. In: Communication papers of the 17th conference on computer science and intelligence systems, pp 3–9. PTI. https://doi.org/10.15439/2022f56
5. Rani S, Ahmad T, Masood S (2023) Handling class imbalance problem using oversampling techniques for breast cancer prediction. Presented at the June 16 (2023). https://doi.org/10.1109/reedcon57544.2023.10150702
6. Cai W, Ning X, Zhou G, Bai X, Jiang Y, Li W, Qian P (2023) A novel hyperspectral image classification model using bole convolution with three-direction attention mechanism: small sample and unbalanced learning. IEEE Trans Geosci Remote Sens, 61. https://doi.org/10.1109/TGRS.2022.3201056
7. Liu J, Guo F, Gao H, Huang Z, Zhang Y, Zhou H (2021) Image classification method on class imbalance datasets using multi-scale CNN and two-stage transfer learning. Neural Comput Appl. https://doi.org/10.1007/s00521-021-06066-8
8. Johnson JM, Khoshgoftaar TM (2019) Survey on deep learning with class imbalance. J Big Data. 6. https://doi.org/10.1186/s40537-019-0192-5
9. Hasanin T, Khoshgoftaar TM, Leevy JL, Bauder RA (2020) Investigating class rarity in big data. J Big Data, 7. https://doi.org/10.1186/s40537-020-00301-0
10. Sadhukhan P, Pakrashi A, Palit S, Mac Namee B (2021) Integrating unsupervised clustering and label-specific oversampling to tackle imbalanced multi-label data
11. Alaba SY, Nabi MM, Shah C, Prior J, Campbell MD, Wallace F, Ball JE, Moorhead R (2022) Class-aware fish species recognition using deep learning for an imbalanced dataset. Sensors, 22. https://doi.org/10.3390/s22218268
12. Guo Y, Fang Z, Zhang S, Dong H (2021) Classification of potato early blight with unbalanced data based on GhostNet. In: 2021 3rd International academic exchange conference on science and technology innovation, IAECST 2021, pp 559–563. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/IAECST54258.2021.9695532
13. Sun M, Jiang A, Li Z (2021) data enhancement for melanoma classification. In: Proceedings—2021 2nd international conference on artificial intelligence and computer engineering, ICAICE 2021, pp 149–155. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICAICE54393.2021.00037
14. Shafay M, Hassan T, Ahmed A, Velayudhan D, Dias J, Werghi N (2022) Programmable broad learning system to detect concealed and imbalanced baggage threats. In: 2022 2nd international conference on digital futures and transformative technologies, ICoDT2 2022. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICoDT255437.2022.9787420

15. Xu X, Li W, Duan Q (2021) Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification. Comput Electron Agric, 180. https://doi.org/10.1016/j.compag.2020.105878
16. Sambasivam G, Opiyo GD (2021) A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks. Egyptian Inf J 22:27–34. https://doi.org/10.1016/j.eij.2020.02.007
17. Airport Security Scanning: Xray Dataset, https://universe.roboflow.com/airport-security-scanning/xray-upzdb. Accessed 27 August 2023
18. Qi S, He M, Hoang R, Zhou Y, Namadi P, Tom B, Sandhu P, Bai Z, Chung F, Ding Z, Anderson J, Roh DM, Huynh V (2023) Salinity modeling using deep learning with data augmentation and transfer learning. Water (Switzerland), 15. https://doi.org/10.3390/w15132482
19. Junaidi A, Tanjung NAF, Wijayanto S, Lasama J, Iskandar AR (2021) Overfitting problem in images classification for egg incubator using convolutional neural network. In: 2021 9th international conference on cyber and IT service management, CITSM 2021. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/CITSM52892.2021.9588815
20. Zhang F, Zhang Z (2023) Information transfer in multitask learning. Data Augment Beyond 37:16220–16221
21. Rihal S, Assal H (2023) Machine learning for the documentation, prediction, and augmentation of heritage structure data. The international archives of the photogrammetry, remote sensing and spatial information sciences. XLVIII-M-2–2023, 1301–1307. https://doi.org/10.5194/isprs-archives-xlviii-m-2-2023-1301-2023
22. Liu X, He J, Liu M, Yin Z, Yin L, Zheng W (2023) A scenario-generic neural machine translation data augmentation method. Electronics (Switzerland), 12. https://doi.org/10.3390/electronics12102320
23. Keller K (2019) Entropy measures for data analysis: theory, algorithms and applications. https://doi.org/10.3390/e21100935.
24. K B (2020) data dimensionality reduction techniques: review. Int J Eng Technol Manag Sci 4:62–65. https://doi.org/10.46647/ijetms.2020.v04i04.010
25. Krizhevsky A, Sutskever I, Hinton GE (2021) Imagenet classification with deep convolutional neural networks
26. Chen F, Tsou JY (2021) DRSNet: Novel architecture for small patch and low-resolution remote sensing image scene classification. Int J Appl Earth Observ Geoinf, 104. https://doi.org/10.1016/j.jag.2021.102577
27. Smith M, Li Z, Landry L, Merz KM, Li P (2023) Consequences of overfitting the van der Waals Radii of ions
28. Baskin C, Liss N, Zheltonozhskii E, Bronshtein AM (2017) Mendelson, a.: streaming architecture for large-scale quantized neural networks on an FPGA-based dataflow platform.https://doi.org/10.1109/IPDPSW.2018.00032
29. Junaidi A, Lasama J, Adhinata FD, Iskandar AR (2021) Image classification for egg incubator using transfer learning of VGG16 and VGG19. In: 10th IEEE international conference on communication, networks and satellite, comnetsat 2021—Proceedings. pp. 324–328. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/COMNETSAT53002.2021.9530826
30. Ruvinga C, Malathi D, Dorathi Jayaseeli JD (2020) Human concentration level recognition based on vgg16 cnn architecture. Int J Adv Sci Technol 29:1364–1373
31. Mungloo-Dilmohamud Z, Khan MHM, Jhumka K, Beedassy BN, Mungloo NZ, Peña-Reyes C (2022) Balancing data through data augmentation improves the generality of transfer learning for diabetic retinopathy classification. Appl Sci (Switzerland) 12. https://doi.org/10.3390/app12115363
32. Zhang Y, Yue J, Song A, Jia S, Li Z (2023) A High-similarity shellfish recognition method based on convolutional neural network. Inf Process Agric 10:149–163. https://doi.org/10.1016/j.inpa.2022.05.009

Mindfulness Intervention Affects Cognitive Abilities of Students: A Time–Frequency Analysis Using EEG

Trupti Taori, Shankar Gupta, Ramchandra Manthalkar, and Suhas Gajre

Abstract The instantaneous frequency measurement is the primary focus of different variably-dimensions signal processing applications. It addresses the non-stationarity of signals globally and the spread of signal frequency locally across time. The present work evaluates the effect of a mindfulness intervention on cognitive workload through different frequency step sizes utilizing one dimensional Gabor function and discrete wavelet transform (DWT) on Electroencephalogram (EEG) signals. Time–frequency tiling is carried out for four different cognitive workload levels of a visual task which shows variation in frequencies along the timeline. The statistical analysis utilizes the mean of small segments (100 ms) from long-duration (7 s) EEG data. The significant difference is found in post-meditation data than pre-meditation data for 90% subjects (18 Subjects out of 20) using a paired t-test with p-value < 0.001 for frontal and occipital electrodes from left and right hemisphere. The right hemisphere shows higher modulation of α activity than the left hemisphere. The results using the Gabor function show good performance by meeting the generalized uncertainty principle. The results with DWT show 'what is where' on time as well as frequency scales clearly with orthogonal basis. Moreover, the cognitive load of four levels is discriminable for an individual with pre-meditation data of twenty subjects, which is clarified with behavioral measures and using statistical analysis on objective measures (p-value < 0.05). The empirical evidence based on EEG signal analysis clarifies that the proposed method is used as an initialization step to quantify cognitive workload based on its emergent frequency at a given time scale for developmental research. This initial investigation holds relevance in determining the optimal duration for implementing meditative programs among students.

Keywords Time–frequency analysis · Cognitive-workload · Mindfulness

T. Taori (✉) · S. Gupta · R. Manthalkar · S. Gajre
Shri Guru Gobind Singhji Institute of Engineering and Technology, Nanded, India
e-mail: trupti.mohota@gmail.com

T. Taori
School of Engineering, Ajeenkya D Y Patil University, Pune, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_15

1 Introduction

The core tasks of the human brain such as thinking, learning, retaining information, and paying attention are governed by cognitive functions. Maintaining a proper level of cognitive function is very important for problem-solving and effective decision-making in the current information-rich world. Multitasking by the brain is well understood through the neuroplasticity that triggers and switches the attention between tasks and improves cognitive construct. Brain training has proved to be an important tool in education, mental health, and development [1–3].

Mindfulness training is the solution to improve the cognitive abilities of the individual (especially students), which helps them in learning and intensify self-consciousness: a key to a happy and successful life, which allow them to stay away from stress and anxiety and improves their attention level [4–6]. Physiological signal analysis is widely used to assess the cognitive load, specifically, the frequency-related signal properties are very important [6–8]. Spectral changes in spontaneous brain activity for various frequency bands are used over a wide range [9]. Coherence analysis is a valuable tool for studying the functional interactions between different brain regions, which assesses the degree to which the signals from two or more EEG electrodes are synchronized at specific frequency bands. The inter brain synchrony in terms of phase locking value is explored in desired frequency band for identification of role of mindfulness during task processing [10]. The STFT provides a time–frequency representation of a signal and allows to see which frequencies are present in the signal at different points in time [10], but with STFT resolution in time and frequency is limited by the choice of window size and overlap. Instantaneous frequency representation uses signal decomposition techniques like wavelet transforms, and wavelet packet transforms [11], but it demands increased computational complexity compared to DWT. However, 1D Gabor function provides precise localization of events in the time–frequency domain through careful tuning of Gabor parameters for optimal results [12]. Thus, DWT and 1D Gabor function is compared in the proposed work for specific measurement of cognitive load using EEG signals as the time frequency analysis is still under exploration in the research community. EEG time series is characterized as a narrow range of locally occurring frequencies, where characteristic frequencies vary across the timeline of the signal. The present work focuses on the analysis of one-dimensional, non-stationary EEG signals for assessing the effect of meditation intervention on cognitive load. The contributions of the proposed work are, (i) Utilization of 1D Gabor function and DWT for decomposition of EEG signal into overlapping and non-overlapping frequency bands respectively. (ii) Utilization of statistical feature (mean) of segmented data (100 ms non-overlap segments) for analysis of pre- and post-meditation data. (iii) Provide relation between cognitive load levels and frequency content across the temporal domain, which is successfully demonstrated using time–frequency tiling of the decomposed EEG data with both decomposition methods. (iv) In addition, confirmation of discriminative load levels of task and improvement in cognitive abilities from pre- to post-meditation is verified using statistical analysis. In the remainder of this paper, Sect. 2 reviews

the existing work related to the cognitive workload through meditation intervention. Section 3 presents the design of visual tasks and data preprocessing. Sections 4 and 5 present the proposed methodology and result. Finally, Sect. 6 concludes the proposed research work.

2 Literature Work

The cognitive workload estimation and classification using physiological signals are carried out extensively in literature [13–16]. Few studies discuss the effect of mindfulness on cognitive load using fNIRS, fMRI, and EEG [8, 17–19]. Gupta et al. show a significant improvement in the cognitive ability of 77.77% of subjects for high complexity levels of visual cognitive load tasks after mindfulness intervention using EEG data [17]. Their work utilizes cognitive ability modulation index, fractal dimension, and state transition through the Hidden Markov model as a measure of estimation. The study in [8], demonstrates a successful transition from an attentive state to a mindful one, as evidenced by changes in alpha brainwave activity during meditation. The strengthened connections in the frontal area with improvement in fractal dimension, highlight the positive effects of mindfulness on brain connectivity and potentially on cognitive processes. Similarly, work in [20] demonstrates modulation in alpha activity during breathing mindfulness exercises. In [21], work demonstrates that the Vipassana meditation influences the quantity and quality of EEG dynamics during cognitive task processing. The study discovered reduced theta synchrony, enhanced alpha de-synchrony, and lesser theta-alpha coherence in proficient meditators. In [22], brief mindfulness training results in the improvement of attention in self-regularized manner is demonstrated through changes in neuronal activity related to attentional control. In [23], the focused attention on breathing technique is evaluated for seven levels of cognitive load on EEG signals using the higher-order crossing features. The work compares reaction time and effect on the brain lobe for pre and post experiment. A study in [24, [25], shows that frequency-specific effects modulate discrete frequency bands and assess the enhancement of the beta and alpha components respectively that affect attentional processing. In [26], authors follow the reverse approach, their work uses quantitative electroencephalogram as a tool to detect weak and strong cognitive skills and uses this screening to adjust meditation practice for an individual that maximizes results of personal wellbeing and minimizes the risk of potential negative effects. In [27], the study verifies the EEG spectral modulations as per the subjective experience of meditation by participant. The alpha frequency/amplitude and 1/f score parameters are evaluated. The study uses breath focus meditations and concludes that meditators have a greater level of focus and less mind wandering. In [28], classification of meditator (Sudarshan Kriya, in combination with yoga) and nonmeditator with statistical features from six frequency bands is carried out using an artificial neural network classifier. In [29], integrated yoga-meditation is explored and demonstrated it as a technique for sustainable relaxation.

The review article [30], discusses the meditation and non-meditation signal classification using machine learning classifiers and articulates the need for further research to identify EEG signatures for meditation intervention. In [10], The mindfulness group exhibited greater theta inter-brain synchrony in the frontal region during the cooperative computer game tasks compared to the non-mindfulness group. In [31], a decrease in alpha-theta cross-frequency coupling is observed for novice individuals undergoing mindfulness training. These findings suggest that EEG parameters related to alpha-theta cross-frequency interactions could serve as effective measures for quantifying and enhancing mindfulness meditation training. Research in [32], concludes that theta1, theta2, alpha1, and alpha2, are associated with distinct cognitive processes. Also, power and coherence analysis can be utilized to effectively classify different meditation practices. Earlier studies have established that a three-month regimen of meditation training resulted in decreases in EEG beta band power during exercises focused on mindfulness of breathing. Additionally, less pronounced decreases are noted in the high alpha frequency range, as well as in the individual peak alpha frequency [9].

3 Material and Methods

The experimentation protocol and preprocessing of the signal are discussed here.

3.1 Experimentation Task

An experiment of a visual task to induce cognitive load of multiple levels is executed two times with a difference of 8 weeks. All twenty participants (students at an engineering institute) with age 20.7 $\pm$ 2.2 years are healthy, right-handed, and do not have any history of psychological disorder. They have the normal or corrected-to-normal vision. The scientific purpose and experiment's nature is informed orally to all participants, and they submitted their written consent for voluntary involvement in the experiment. The institutional ethical committee has approved the experiment design. The visual task uses the visual protocol which emphasize on the recognition and counting of the basic trigonometric shapes to induce cognitive workload of different four levels (L1, L2, L3, L4). At L1 level two different shapes are shown, at L2 level three different shapes are shown, at L3 level four different shapes are shown where one of the shapes is inscribed in other shape, and at L4 level four different shapes where two shapes inscribed in other shapes are shown. From L1 to L4, levels are designed with increasing complexity. The sample trial of L1 and L4 level are shown in Fig. 1a, b respectively.

Every individual level consists of ten trials. In the timeline of an individual trial, the visual stimulus period (7 s) displays an image on the screen, then the screen goes blank (2 s), which is followed by a maximum response time (5 s) duration. It is

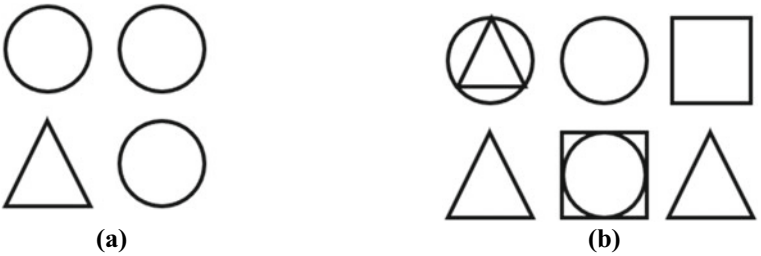


Fig. 1 **a** Sample trial of level L1 **b** Sample trial of level L4

expected that the participants must remember the count of the trigonometric shapes and provide the correct answer to an asked question. During the response period, the participants must provide the answer by pressing a correct option key after reading a question. There is a 30 s relaxation period before starting the very first trial of each level. Subjective rating, NASA task load index is obtained from all participants after completion of each level. EEG signals are captured when participants are involved in the execution of the said visual task, before starting the meditation protocol, referred to as pre-meditation data. In mindfulness training, all the novice participants practice focused attention for 20 min a day for continuous eight weeks, where participants must concentrate on incoming and outgoing breaths. Again, after eight weeks, the same visual experiment is executed, and EEG data is collected referred to as post-meditation data. 32-channel wireless ENOBIO (Neuro electric, Spain) is used for EEG data collection at a sampling rate of 500 Hz. Dry electrodes are placed according to the 10–20 international system on the scalp, and the right mastoid is used as a reference. The more experimentation details can be referred from [33].

3.2 Data Preprocessing

Before analysis, the extraneous information from the raw EEG data is removed. The DC offset removal is obtained by subtracting the average value from the entire data channel. For noise elimination, Hamming windowed FIR filter having a pass-band of 0.1 to 45 Hz is used. Artifact identification and removal are done through ADJUST1.1.1 (plug-in in EEGLAB), a fully automated algorithm that removes the independent component from data. ADJUST is optimized to capture eye movements, blinks, and generic discontinuities using a combination of spatial and temporal data features.

4 Methodology

This section discusses the details of signal decomposition using 1D Gabor transform and DWT, Time segmentation, and features used in the analysis.

4.1 Signal Decomposition

Time–frequency transform are very important for discovering the pattern of frequency changes in signal and to isolate the distinct components from signal for further analysis. The proposed work uses two different methods for signal decomposition: one dimensional Gabor function, and DWT.

(1) *Decomposition using 1D Gabor Function:* The one-dimensional Gabor function is defined in terms of its even and odd part using the following equation as,

$$g_e(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-t^2}{2\sigma^2}} \cos(2\pi\omega_0 t) \tag{1}$$

$$g_o(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-t^2}{2\sigma^2}} \sin(2\pi\omega_0 t) \tag{2}$$

where sine/cosine wave is multiplied with a Gaussian window. ω_0 and σ represents the centre frequency and the spread of the Gaussian window, respectively. The Fourier transform of the Gabor function is given as follows,

$$\hat{g}(\omega) = \sqrt{2\pi}\sigma e^{-j\omega t_o} e^{\frac{-\sigma^2\omega^2}{2}} \tag{3}$$

The time dispersion and spectral bandwidth of Gabor function and Gaussian function are same and thus it minimizes their product. Therefore, for signal processing the minimum in the uncertainty principle is obtained using the Gabor function. The product of the spread of a signal in the time domain (Δt) and frequency domain (Δf), as per uncertainty principle, is greater or equal to a fixed constant (C). From Eqs. 1–3 it is clear that these family of functions cover (t, ω) space completely. In the proposed work Gabor frames are formed with four different center frequencies (14, 20, 26, 32 Hz) and σ is computed as,

$$\sigma = \frac{11}{2\pi\omega_0} \tag{4}$$

Here, the numerator indicates the number of cycles of the Gabor function per standard deviation of the envelope. The convolution operation is important in Time–frequency

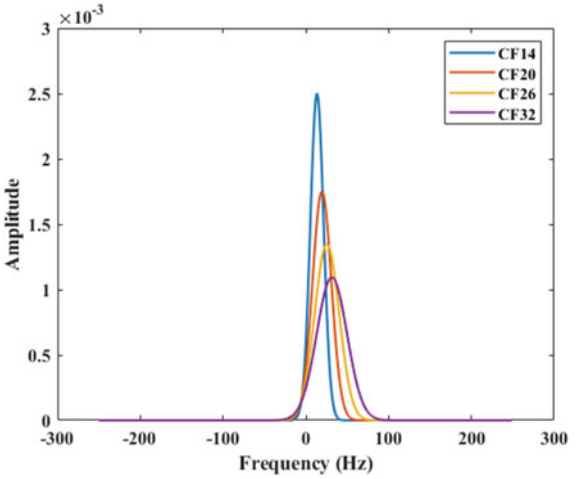


Fig. 2 Frequency bands representation using Gabor frames with center frequency shown in the legend

analysis and is carried out by taking the dot product of two signals elementwise and then summing them together. In the proposed work, filtered EEG data is convolved with only the even component (cos) of the Gabor frames to decompose data in frequency bands, where the frequency range of interest is from 8 to 46 Hz. The 3 dB down frequency range for respective center frequencies are used. These overlapped frequency bands are depicted in Fig. 2.

(2) *Decomposition using DWT:* The spatio-frequency space is also spanned with scaling and translation of basic function in wavelet. Wavelet is one more method other than the Gabor function for time–frequency localization. DWT is preferred over Continuous wavelet transform for various practical applications. Conventionally the scaling parameter takes 2^l values and thus creates a dyadic wavelet. In the present work, the frequency of interest is separated into four frequency bands (4–8 Hz, 8–16 Hz, 16–32 Hz, 32–64 Hz) using DWT (biorthogonal wavelet, 'bior 4.4').

4.2 Time Segmentation and Feature Used

Information extraction and component identification is a key to analyzing a long EEG time series [34]. In visual aspects, a 100–300 ms window is sufficient for identification and decoding after perception [35], thus in the proposed work decomposed EEG data is segmented in 100 ms non-overlap segments. Further, statistical feature, mean is extracted for these segments. The mean is computed as the average of the total N signal data points for time series X = × 1, × 2, ……xN.

$$M = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{5}$$

Table 1 Details of behavioral measures: response time, accuracy, and NASA TLX score

Levels	Response time	Accuracy of answers	NASA TLX score
L1	3.33 ± 0.47	99.31 ± 2.54	45.22 ± 4.06
L2	3.52 ± 0.39	99.01 ± 1.78	54.51 ± 5.01
L3	3.91 ± 0.27	97.64 ± 1.94	66.24 ± 2.04
L4	4.33 ± 0.28	96.59 ± 4.79	82.13 ± 6.25

5 Results

5.1 Behavioral Analysis

The outcome of behavioral measure to discriminate task levels is mentioned in Table 1. For pre-meditation data, behavioral measures show that the response time for answers increases, the accuracy of giving correct answers decreases, and the NASA TLX score increases from L1 to L4. These results reveal that the complexity of task levels increases from L1 to L4 which induces incremental level of visual cognitive load.

5.2 Subject Specific Analysis

The objective measures consider pre- and post-meditation EEG data of 20 subjects. The existing studies claim that the frontal and occipital lobes are related to working and visual perception, respectively [11, 36, 37]. As the proposed work utilizes visual stimulus to induce cognitive workload, the data of eight electrodes from frontal region, and two electrodes from occipital region (FP1, FP2, AF3, AF4, F3, F4, F7, F8, O1, O2) out of 32 available electrodes are used throughout the analysis. Non-stationary EEG data is subjected to inter-subject and intra-subject variability, due to variation in data across sessions and subjects. Thus, the subject-specific evaluation is carried out in the proposed study. All ten trials from every individual level are averaged together. From visual stimulus period, 7 s data (3500 data points) is only considered for all the trials from L1 to L4. From total relaxation period of 30 s, only 7 s data exactly before starting of the first trial of every level is considered and is referred to as R1 to R4. Further, the data is segmented into 70 non-overlap segments of 100 ms each (50 data points). The mean of the segmented data is calculated for an individual subject. Further, the mean of respective relaxation period and respective level trials are subtracted (L1-R1, L2-R2, L3-R3, L4-R4) for pre- as well as post-meditation data. With subtraction of the mean values of the signal and baseline, the data is effectively normalized, which removes any overall offset or bias in the data. This is crucial for accurate comparison of the relative changes or differences between the two. The square of this difference is obtained which shows the variation in its

amplitude on a time scale across all levels. Squaring the differences can help capture the variability and meet assumptions of normality or homogeneity of variances.

The paired t-test is applied on this squared difference for every two successive levels of pre-meditation data. The result shows the significant difference (p-value < 0.05) between levels of a designed task, indicates that all four levels are discriminable based on cognitive load induced with visual protocol. The squared difference of pre- and post-meditation data for respective levels is also compared using paired t-test for all frequency bands of individual level. Eighteen subjects out of total 20 subjects show significant difference between pre- and post-meditation data with p-value < 0.001. A time–frequency tiling for all data points with four frequency bands of F3 electrode for pre- and post-meditation 1D Gabor decomposed and DWT decomposed data of two different subjects are depicted in Figs. 3 and 4, respectively. The pre-meditation data shows higher amplitude than post-meditation data across all levels in 1D Gabor decomposed data for subject two, whereas the pre-meditation data shows less amplitude than post-meditation data across all levels in DWT decomposed data for subject three. The effects of meditation on EEG can be context-dependent, they may differ based on factors such as the specific meditation technique and the individual's level of expertise. From Fig. 3, it is clear that with increasing complexity level of the task the dominance of frequencies is from higher to lower in pre-meditation 1D Gabor decomposed data. Whereas for DWT decomposed data, lower frequencies are dominant across levels of a task. The variations of frequency surrounding the center frequencies along the time scale are visible across all levels of the task in Fig. 3, whereas Fig. 4 shows the particular frequency at each time point for DWT.

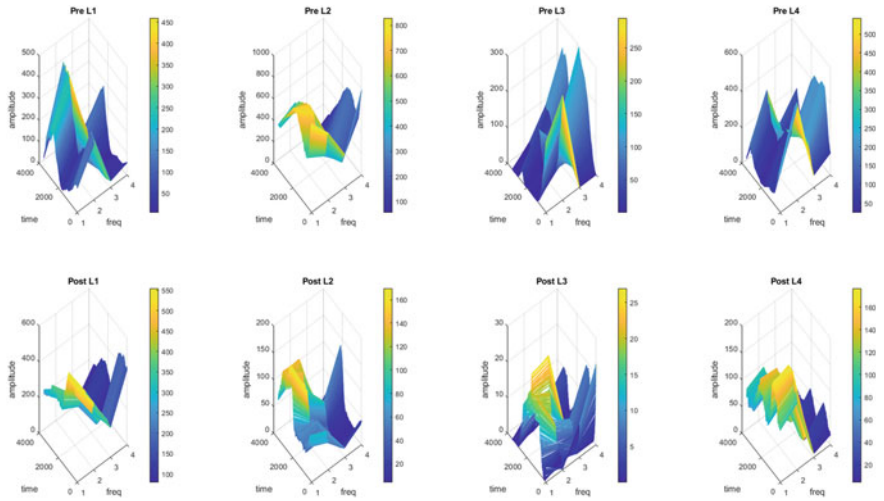


Fig. 3 Time–frequency tiling for Pre- and Post-meditation 1D Gabor decomposed data of L1-L4 for F3 electrode of S2

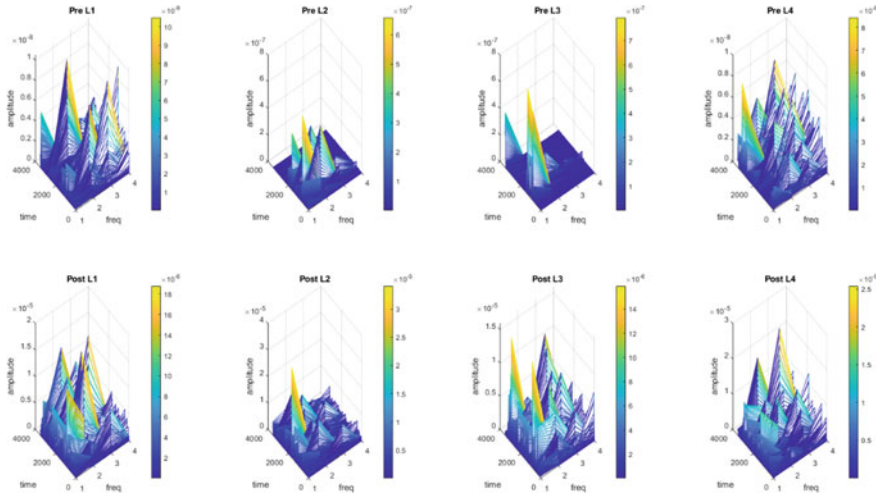


Fig. 4 Time–frequency tiling for Pre- and Post-meditation DWT decomposed data of L1-L4 for F3 electrode of S3

(a) *Left and Right Lobe Analysis*

In the present study, the left and right cerebral hemispheres are compared for visual task and this analysis reveals that there is a significant difference in frequencies during activity performance. The ratio of electrode pairs (mentioned in Table 2) from left and right hemispheres are computed for pre- and post-meditation data. Out of 20 subjects 18 subjects show the significant difference in left and right cerebral hemisphere activity (p-value < 0.001) indicated for single subject in bold face in Table 2. The frontal region electrodes show modulations in amplitude of θ, α, and β frequency bands. The right hemisphere shows more dominance than the left hemisphere which is matching with the existing study reported in [19]. Later these ratios are compared using paired t-test, the calculated p-value for all levels and two bands (α and β) with nearest same frequency in 1D Gabor and DWT decomposition are mentioned in Table 2 for single subject. Here 1D Gabor decomposed data shows significant difference in both the frequency bands 1 and 2 (with center frequency 14 and 20 Hz) for all task levels, whereas the DWT decomposed data shows significant difference in only one band 2 (8–16 Hz). It could be attributed to the frequency leakage in Gabor decomposition and its overlapping band structure as the basis created by the Gabor function is not orthogonal, where as the basis designed for the wavelet is orthogonal, and no frequency overlapping is seen among bands in DWT. This result is also aligned with the existing study [20] which concludes that α activity power increases during mindfulness.

(B) Comparison with Existing Work

A comparison of the present study with existing studies is presented in Table 3. Fair comparison is not possible as studies use different types of meditation with variable duration and again the method of analysis is also varying.

Table 2 Statistical paired t-test between pre- and post- meditation data of left and right electrode pair ratio for S1. Bold face p-values represent significant difference

Electrode ratio	Task level	Paired T-test between electrode ratios of pre- and post-meditation data (1D Gabor decomposed data)		Paired T-test between electrode ratios of pre- and post-meditation data (DWT decomposed data)	
		Freq-Band1	Freq-Band2	Freq-Band2	Freq-Band3
FP1/FP2	L1	**2.7902e-12**	0.3122	**4.2918e-30**	0.3412
	L2	0.3032	**1.2735e-13**	**1.9470e-37**	0.1812
	L3	**1.9773e-15**	0.1248	**4.2556e-77**	0.3833
	L4	**5.7796e-23**	0.0013	**1.4403e-69**	0.8215
AF3/AF4	L1	**8.3737e-06**	**1.3090e-16**	**3.4124e-08**	0.4519
	L2	**3.7455e-24**	**7.1039e-18**	**1.4756e-18**	0.4443
	L3	**1.0259e-26**	**6.8758e-21**	**2.9579e-09**	0.3111
	L4	**7.7531e-14**	**1.0855e-24**	**1.6107e-24**	0.3256
F3/F4	L1	0.9443	**6.4859e-27**	**1.0108e-87**	**2.3103e-07**
	L2	**7.0168e-25**	0.0146	**2.5148e-74**	**2.6034e-04**
	L3	**6.6587e-28**	**3.7029e-25**	**4.8031e-41**	**1.9980e-09**
	L4	**1.3155e-34**	**6.3813e-26**	**4.6760e-18**	0.3164
F7/F8	L1	**1.4307e-04**	**9.3349e-27**	**9.9020e-49**	0.5907
	L2	**1.3619e-26**	**9.3642e-15**	**4.9040e-68**	0.2857
	L3	**2.3897e-29**	**2.8102e-35**	**1.0864e-68**	0.3165
	L4	**1.3896e-35**	**6.0589e-29**	0.2754	0.3189
O1/O2	L1	**3.3424e-08**	0.2510	**2.1300e-28**	0.2976
	L2	**4.0614e-19**	**2.7365e-22**	**1.9814e-61**	0.4774
	L3	**4.2379e-32**	0.2605	0.1296	0.2274
	L4	**1.3530e-37**	**8.3192e-05**	0.0802	0.1788

Table 3 Comparison of the proposed method with existing work

Author	Meditation Type and Duration	Analysis Method	Remark
Meshram et al. [38]	Sudarshan Kriya Yoga	Time–Frequency analysis through wavelet decomposition using Coiflet	The Wavelet Transform emerges as the most effective technique for conducting time–frequency analysis on EEG signals. This method not only furnishes essential frequency data but also pinpoints the precise time moments of occurrence
Aftanas et al. [39]	Sahaja Yoga Meditation	Band power obtained using FFT	The narrow frequency bands of theta and alpha activities serve as indicators of the specialized neuronal networks engaged in various cognitive and emotional processes
Bing-Canar et al. [20]	Mindful-breathing exercise	Alpha power and error-related alpha suppression	Individuals engaged in the mindful-breathing condition exhibited amplified alpha power while participating in the listening exercise, followed by improved Event-Related Alpha Synchronization during the Stroop task
Gupta et al. [17]	focused attention meditation	Cognitive ability modulation index, fractal dimension and state transition using the Hidden Markov model	Improvement in CAMI and decrement in FD is seen after mindfulness practice 77.77% subjects perform better after mindfulness training for visual task
Proposed method	Focused attention meditation	Time -frequency analysis, employing DWT and 1d Gabor Function	The significant difference is found in post-meditation data than pre-meditation data for 90% subjects (18 Subjects out of 20) for frontal and occipital electrodes from left and right hemisphere. The right hemisphere shows higher modulation of α activity than the left hemisphere

6 Conclusion

The proposed method uses time–frequency analysis utilizing the 1D Gabor function and DWT decomposition on EEG data. The difference in the mean of multiple levels reveals that the frequency-specific changes are occurring in the non-stationary signal which can be verified in the time–frequency tiling of data using both decomposition methods. The cognitive load levels L1-L4 are discriminable (p-value < 0.05) and there is a significant difference between pre- and post-meditation data is justified

with the statistical paired t-test (p-value < 0.001), considering electrodes of frontal and occipital lobes from left and right hemispheres. The alpha activity shows significant modulation from pre- to post-meditation, which shows improvement in cognitive abilities of subjects. This initial study carries implications for determining the optimal duration of meditative programs for students. The study suggests that meditation practice can influence EEG signals, it's not accurate to make a blanket statement that meditation always leads to higher signal amplitude. The relationship between meditation and EEG activity is complex and multifaceted, and further research is needed to fully understand the nuances of this interaction. The time frequency technique can be further extended to quantify multiple cognitive load levels and their exact related emergent frequencies across the timeline of a non-stationary EEG signal. With smaller segment duration non-stationary issues can be handled efficiently and the proposed technique shows potential that it can also be used in future for micro-state analysis which involves segmenting EEG data into short, quasi-stable periods of topographical configurations in understanding the synchronous mechanism of the brain. Time–frequency analysis complements this by providing information on the oscillatory dynamics within these micro-states.

References

1. Slagter HA, Davidson RJ, Lutz A (2011) Mental training as a tool in the neuroscientific study of brain and cognitive plasticity. Front Hum Neurosci 5:17
2. Doborjeh Z, Doborjeh M, Taylor T, Kasabov N, Wang GY, Siegert R, Sumich A (2019) Spiking neural network modelling approach reveals how mindfulness training rewires the brain. Sci Rep 9(1):6367
3. Gao J, Sun R, Leung HK, Roberts A, Wu BWY, Tsang EW, Sik HH (2023) Increased neuro-cardiological interplay after mindfulness meditation: a brain oscillation-based approach. Front Hum Neurosci 17:1008490
4. Ng HYH, Wu CW, Huang FY, Huang CM, Hsu CF, Chao YP, ... Chuang CH (2023) Enhanced electroencephalography effective connectivity in frontal low-gamma band correlates of emotional regulation after mindfulness training. J Neurosci Res 101(6):901–915
5. Naderan M, Ghoshuni M, Afrouz EP (2021) The effect of mindfulness training on creativity in healthy subjects: a pilot EEG study. Pol Psychol Bull 52(4):327–333
6. Shanok NA, Saldias-Manieu C, Mize KD, Chassin V, Jones NA (2023) Mindfulness-training in preadolescents in school: The role of emotionality, EEG in theta/beta bands, creativity and attention. Child Psychiatry Hum Dev 54(4):1152–1166
7. Li J et al (2016) Decoding EEG in cognitive tasks with time-frequency and connectivity masks. IEEE Trans Cognit Dev Syst 8(4):298–308
8. Do H, Hoang H, Nguyen N, An A, Chau H, Khuu Q, ... Ha H, Intermediate effects of mindfulness practice on the brain activity of college students: an EEG study. IBRO Neurosci Reports 14:308–319
9. Skwara AC, King BG, Zanesco AP, Saron CD (2022) Shifting baselines: longitudinal reductions in EEG beta band power characterize resting brain activity with intensive meditation. Mindfulness 13(10):2488–2506
10. Deng X, Yang M, Chen X, Zhan Y (2023) The role of mindfulness on theta inter-brain synchrony during cooperation feedback processing: an EEG-based hyperscanning study. Int J Clin Health Psychol 23(4):100396

11. Rajendran VG, Jayalalitha S, Adalarasu K (2021) EEG based evaluation of examination stress and test anxiety among college students. IRBM
12. Kumar TS, Kanhangad V, Pachori RB (2015) Classification of seizure and seizure-free EEG signals using local binary patterns. Biomed Signal Process Control 15:33–40
13. Iqbal MU, Srinivasan B, Srinivasan R (2020) Dynamic assessment of control room operator's cognitive workload using Electroencephalography (EEG). Comput Chem Eng 141:106726
14. Gupta SS et al (2021) Classification of cross task cognitive workload using a deep recurrent network with modeling of temporal dynamics. Biomed Signal Proc Control 70:103070
15. Taori TJ, Gupta SS, Gajre SS, Manthalkar RR (2022) Cognitive workload classification: towards generalization through innovative pipeline interface using HMM. Biomed Signal Process Control 78:104010
16. Taori T et al (2022) Cross-task cognitive load classification with identity mapping-based distributed CNN and attention-based RNN using Gabor decomposed data images. IETE J Res, 1–17
17. Gupta SS, Manthalkar RR, Gajre SS (2021) Mindfulness intervention for improving cognitive abilities using EEG signal. Biomed Signal Process Control 70:103072
18. Eby, Lillian T et al (2019) Mindfulness-based training interventions for employees: a qualitative review of the literature. Human Resource Manag Rev 29(2):156–178
19. Pandey P, Miyapuram KP (2021) Nonlinear EEG analysis of mindfulness training using interpretable machine learning. In: 2021 IEEE International conference on bioinformatics and biomedicine (BIBM), pp 3051–3057. IEEE
20. Bing-Canar H, Pizzuto J, Compton RJ (2016) Mindfulness-of-breathing exercise modulates EEG alpha activity during cognitive performance. Psychophysiology 53(9):1366–1376
21. Kakumanu RJ et al (2019) State-trait influences of Vipassana meditation practice on P3 EEG dynamics. Progress Brain Res 244:115–136
22. Moore AW et al (2012) Regular, brief mindfulness meditation practice improves electrophysiological markers of attentional control. Front Human Neurosci 6:18
23. Jadhav N, Manthalkar R, Joshi Y (2016) Analysis of the effect of meditation on cognitive load using higher-order crossing features. International conference on communication and signal processing 2016 (ICCASP 2016). Atlantis Press
24. Egner T, Gruzelier JH (2004) EEG biofeedback of low beta band components: frequency-specific effects on variables of attention and event-related brain potentials. Clin Neurophysiol 115(1):131–139
25. Chow T, Javan T, Ros T, Frewen P (2017) EEG dynamics of mindfulness meditation versus alpha neurofeedback: a sham-controlled study. Mindfulness 8:572–584
26. Fingelkurts AA, Fingelkurts AA, KallioTamminen T (2015) EEG-guided meditation: a personalized approach. J Physiol Paris 109(4–6):180–190
27. Rodriguez-Larios J, Bracho Montes de Oca EA, Alaerts K (2021) The EEG spectral properties of meditation and mind wandering differ between experienced meditators and novices. NeuroImage 245:118669
28. Sharma H, Raj R, Juneja M (2019) EEG signal-based classification before and after combined Yoga and Sudarshan Kriya. Neurosci Lett 707:134300
29. Radhamani R et al (2020) Computational analysis of cortical EEG biosignals and neural dynamics underlying an integrated mind-body relaxation technique. Procedia Comput Sci 171:341–349
30. Kora P et al (2021) EEG based interpretation of human brain activity during yoga and meditation using machine learning: a systematic review. Complementary Therapies Clin Pract 43:101329
31. Rodriguez-Larios J, Wong KF, Lim J, Alaerts K (2020) Mindfulness training is associated with changes in alpha-theta cross-frequency dynamics during meditation. Mindfulness 11:2695–2704
32. Travis F (2020) Temporal and spatial characteristics of meditation EEG. Psychol Trauma Theory Res Pract Policy 12(2):111
33. Gupta SS, Manthalkar RR (2020) Classification of visual cognitive workload using analytic wavelet transform. Biomed Signal Process Control 61:101961

34. Wang Q, Sabrina O (2013) Real-time mental arithmetic task recognition from EEG signals. IEEE Trans Neural Syst Rehabil Eng 21(2):225–232
35. Fabre-Thorpe M (2011) The characteristics and limits of rapid visual categorization. Front Psychol 2:243
36. Keskin M et al (2019) EEG and eye-tracking user experiments for spatial memory task on maps. ISPRS Int J Geo-Inf 8(12):546
37. Chen C-S et al (2020) Prefrontal brain electrical activity and cognitive load analysis using a non-linear and non-stationary approach. IEEE Access 8:211115–211124
38. Meshram Y, Fulpatil P (2012) Review paper on electroencephalographic evaluation of Sudarshan Kriya. Int J Sci Res (IJSR) ISSN (Online), 2319–7064
39. Aftanas LI, Golosheikin SA (2003) Changes in cortical activity in altered states of consciousness: the study of meditation by high-resolution EEG. Hum Physiol 29:143–151

Wide Communication Coverage SpO_2 Monitoring Using Local Host HTML Web Page



I. Dewa Gede Hari Wisana, Nabila Surayya Saidah, Priyambada Cahya Nugraha, Moch Prastawa Assalim Tetra, Dessy Tri Wulandari, and Tetrik Fa'altin

Abstract SpO_2 monitoring needs to be done routinely and can be done anywhere to prevent delays in handling. However, during the COVID-19 pandemic, sterile and limited patient handling is needed to prevent the spread of this disease, so monitoring of patients cannot be carried out optimally and can cause delays in treatment. One of the efforts that can be made to improve service to patients is by monitoring SpO_2 which is carried out based on IoT, one of which is through an HTML Web Page. The purpose of this study is to analyze SpO_2 data sent and received by IoT media. The contribution of this research is knowing the response of SpO2 data sent and received through IoT media for the development of a SpO_2 monitoring system. The procedure for achieving this goal is the MAX30100 sensor whose output will be processed and displayed on an HTML Web Page. This study obtained an average delay time of 219.3 s. In this study, it was concluded that the transmission of SpO_2 data on HTML Web Pages was said to be good. This research is expected to be useful for the development of a SpO_2 monitoring system so that it can improve services for patients, especially those who need sterile treatment.

Keywords MAX30100 · SpO_2 · HTML · Time delay

1 Introduction

Oxygen saturation (SpO_2) is one of the vital parameters recorded through an important vital sign monitor of the human body [1, 2]. SpO_2 is very important for monitoring, because this value affects the condition of vital organs in the body in addition to heart rate, blood pressure, respiratory rate, and temperature [3, 4]. SpO_2 generally

I. D. G. H. Wisana · N. S. Saidah · P. C. Nugraha (✉) · M. P. A. Tetra · D. T. Wulandari · T. Fa'altin
Department of Electromedical Engineering, Health Polytechnic Ministry of Health Surabaya, Surabaya, Indonesia
e-mail: pcn1967@poltekkesdepkes-sby.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_16

depends on the concentration of hemoglobin in red blood cells, which is the normal value in adults in general is 95% to 100% [5, 6]. Vital organs in the body can suffer permanent damage if they lack oxygen supply. Low oxygen levels in the blood can be a major cause of morbidity and death [7, 8]. SpO_2 monitoring is increasingly needed, especially during the COVID-19 pandemic, when there were 35 million people affected in 213 countries. Patients infected with the outbreak had an average SpO_2 level of $\leq$92% [9, 10]. SpO_2 level status is very important to know not only for diseases such as COVID-19 but also for serious diseases related to the human circulatory and respiratory systems. The SpO_2 signal can also be used as an alarm validation of arrhythmias [11, 12]. With the consideration of health information, it is felt necessary to have a real-time monitoring system of the body's vital condition. SpO_2 monitoring needs to be done regularly and can be done anywhere to prevent delays in handling that can cause other serious diseases such as silent hypoxemia [13, 14]. However, during the COVID-19 pandemic, sterile and limited patient handling is required to prevent the spread of this disease, so patient monitoring cannot be done optimally and can cause delays in handling [15, 16]. One of the efforts that can be made to improve services to patients is the monitoring of SpO_2 which is carried out based on IoT (Internet of things). This IoT system allows users and other systems to connect and communicate so that it can be a solution to the problem [17].

In 2018, Lanny Agustine et al., conducted a study about monitoring heart rate for arrhytmia using pulse oximeter using android. This study aims to design an arrhythmia monitoring tool using a pulse oximeter. This tool displays SpO_2 value and heart rate through the Blynk application on the smartphone. However, this tool still has the disadvantage that it only displays SpO_2 and BPM values and has not displayed the form of SpO_2 signals [18]. Then, in 2022, Sari L. et al., conducted a research about monitoring vital sign with BPM and SpO_2 parameter using Tinger.io Platform and notification using Telegram Aplication. This study aims to design BPM and SpO_2 monitoring on the Tinger.io platform which will later send notifications on telegrams if there are certain conditions. This study used the MAX30100 sensor as a sensor for SpO_2 readings. However, this study still displays values only, still does not display SpO_2 signals [19]. In 2021, G. Sidharta et al., conducted a research about pulse oximeter using Internet of Things. This study aims to design a Wi-Fi-based BPM and SpO_2 monitoring tool displayed on the Blynk platform on smartphones. However, in this study there were only BPM values, SpO2, and graphs of these values without SpO_2 signals [20]. From some of the studies above, it can be said that monitoring SpO_2 with a signal display is still a little done. In 2021, Preetham Gade et al., conducted a study about IoT based pulse oximeter using ESP8266. This research aims to design a pulse oximeter that can be monitored via IoT, namely Web Server. In this study using ESP8266 for processing and sending data to a web server, and it is proven that ESP8266 is able to send data on the desired IoT platform [21]. In 2020, O. Y. Tham et al., conducted a study about health monitoring device using IoT with oxygen saturation (SpO_2) and heart rate level as a parameter on the module. This research aims to build a SpO_2 monitoring system and heart rate through IoT. This research uses the MAX30100 sensor for recording SpO_2 data which is then displayed in HTML with firebase as a system database so that it can be accessed

periodically and even real-time. However, this study still only took SpO_2 and BPM values without the display of signals from SpO_2 or ECG [22].

Therefore, from the various studies above, it can be concluded that little is the SpO_2 monitoring system that displays values and signals. In fact, as we know, the value and signal of SpO_2 carry important information about the state of the human body. These studies mostly used existing Web sites and mostly submitted value only. So, in this research we will try to design our own website for monitoring SpO_2 data using HTML Web Pages in the hope of displaying the value and signal of SpO_2 in real time. In addition, based on some of the studies above, the MAX30100 sensor can be used as a sensor module to record SpO_2 in humans, so this sensor will also be used in this study.

Based on the above problems, the author wants to develop research on SpO_2 Monitoring Via HTML Web Page with SpO_2 parameters that will be displayed on the IoT platform, namely HTML Web Page and using the MAX30100 sensor with the Wemos Mega 2560 microcontroller. This study aims to analyze IoT-based SpO_2 signal transmission to facilitate remote monitoring and improve service to patients.

2 Materials and Method

In conducting this research, the researchers conducted it at the Electronics Laboratory in the Department of Electromedical Technology of the Surabaya Ministry of Health Polytechnic and used a calibrated Pulse Oximeter as a comparison tool. If it is in accordance with the comparison, the module testing will be carried out on 10 male respondents with an age range of 20–30 years. Data collection will be carried out for 15 s in a calm state of respondents. The recording of data sent is done through Delphi 7, while the recording of data received through HTML Web Page [23] will then be compared to find time delay.

In this study, the next stage after the design of the tool was completed, the device was tested on a calibrated pulse oximeter to see if the reading value was compatible with the comparison. If the results obtained are correct, data collection is carried out on 10 respondents. The module will record the value and signal from the respondent to be displayed on the TFT LCD and sent on the HTML Web Page. Data retrieval on the module will use the Delphi 7 application to store data in excel form. While data retrieval on HTML Web Page will be done via the web to store data in excel form. Then, the data sent from the module will be compared with the data received by the HTML Web Page to then be analyzed based on time delay.

The independent variable of this study was SpO2 data before it was sent. The bound variable is the SpO2 data after it is sent. And the controlled variable in this study is the Wemos Mega 2560 microcontroller.

Description:

———: Not done by researchers

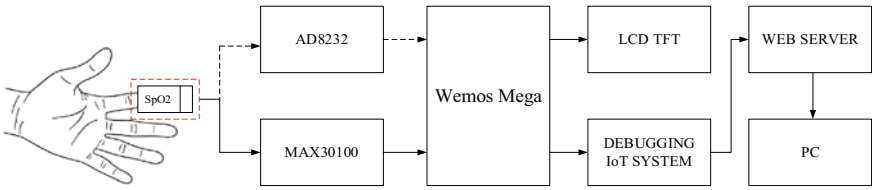


Fig. 1 Block diagram of the system from the SpO2 monitoring module Via HTML

———-: Done by researchers

This study used the MAX30100 sensor as an SpO2 sensor with reflectance type to record SpO2 from respondents [24, 25]. The microcontroller used to process data is Wemos Mega 2560. For the display of the tool using LCD TFT Nextion 5 Inch Enhanced Models [26].

This paragraph will explain Fig. 1 which is a block diagram on the system. The Block Diagram below has 3 main parts to explain how the system works, namely input, process and output. The input consists of an MAX30100 sensor which will later record SpO2 data from the patient placed on the patient's finger. The input from the sensor will be read by the Wemos Mega 2560 microcontroller. The process consists of a Wemos Mega 2560 microcontroller which functions as a data processor from the sensor output and sends data in the form of signals and SpO2 values. The Wemos Mega 2560 microcontroller will send signal data and SpO2 values to be displayed on the TFT LCD as a display on the device as well as the IoT platform which will later be stored on a PC.

While Fig. 2 is a flow diagram of how the system works. Turn on the ON button, after the module is activated, the process begins with initialization after completion, it will continue to the input reading section of the MAX30100 sensor which will then continue to the SpO2 signal detection process and SpO2 value. Data in the form of signals and SpO2 values will be displayed on the TFT LCD. Then, if it is not connected to Wi-Fi then the process will return to displaying data on the TFT LCD. If the device is connected to Wi-Fi, the Wemos Mega 2560 microcontroller will send a signal and SpO2 value on the IoT platform for display. After that, then the system will be completed.

2.1 Data Analysis

The average is a number obtained from dividing the number of data values by the amount of data in the set. This formula will be used when retrieving SpO2 data is carried out 5 times to find the average value of the recorded SpO 2. The average value formula is in Eq. (1):

$$\overline{x} = \frac{x1 + x2 \cdots + xn}{n} \tag{1}$$

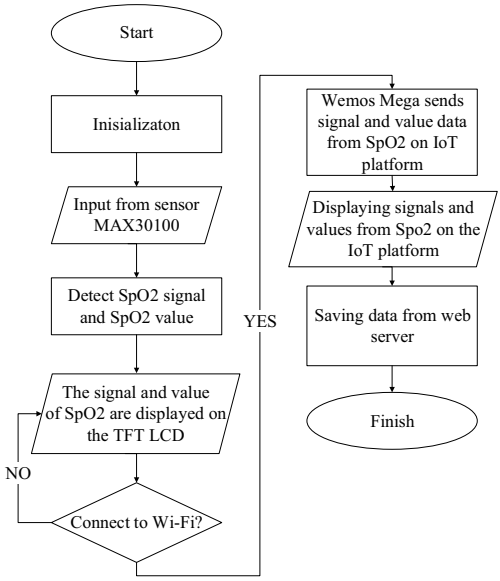


Fig. 2 The system flowchart from the SpO2 monitoring module via HTML web page

where x denotes the mean (mean) for the n-measurements, $\times$ *1* denotes the first measurement, $\times$ *2* denotes the second measurement, *xn* denotes n measurements and % error indicates a system error. The lower error value is the mean difference of each data. Errors may indicate a deviation between the standard and the design or model. The error formula is shown in the Eq. (2):

$$\%\text{error} = \frac{(\boldsymbol{x_n} - \mathbf{x})}{\boldsymbol{x_n}} \times 100\% \tag{2}$$

where *xn* is the measured value of the machine calibrator. X is the measured value of the design.

3 Result

In this study, module testing was carried out with a comparison in the form of a Pulse Oximeter and testing on 10 respondents. Figure 3 is the result of the design of this research module:

The circuit on the module uses the MAX30100 sensor circuit to record SpO2. Output from the sensor will be processed on the Wemos Mega 2560 microcontroller. In addition to processing sensor output, the Wemos Mega 2560 is used to input web design, send data to the web, and store data from the web. After processing, the output will be displayed on the display and HTML web page. The display used in

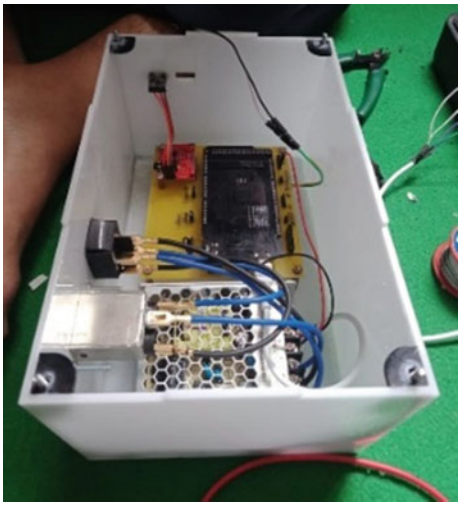


Fig. 3 Circuit results on the SpO2 monitoring module

the module is the Nextion 5 Inch Enhanced TFT LCD. The data displayed is in the form of SpO2 signals, ECG signals, SpO2 values, and BPM as seen in Figs. 4 and 5.

In Fig. 4 shows the web interface used in the monitoring process. With internet access, users can easily access the ECG and BPM signal display from this module, either via smartphone or computer. To begin monitoring, users need to enter the appropriate IP address on their web browser, the IP address entered must match the IP address that appears on the Nextion 5 Inch TFT LCD display. It allows users to monitor heart conditions and measure their heart easily and accurately.

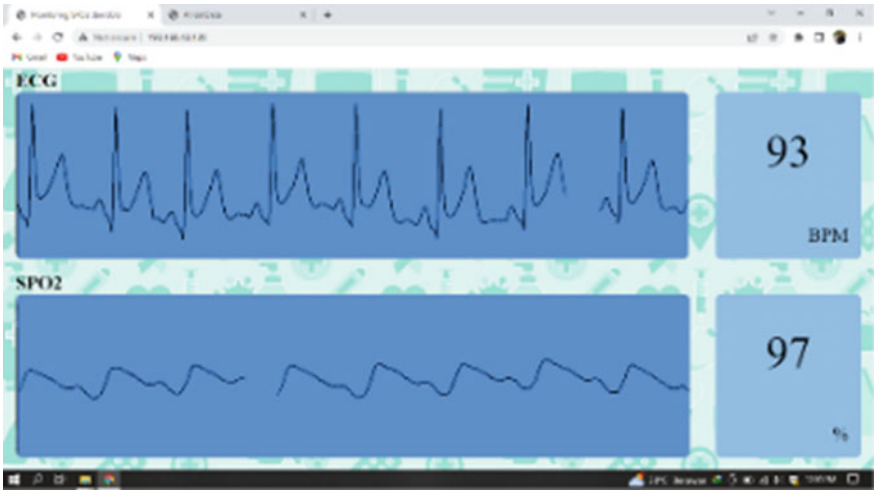


Fig. 4 HTML web page display when monitoring

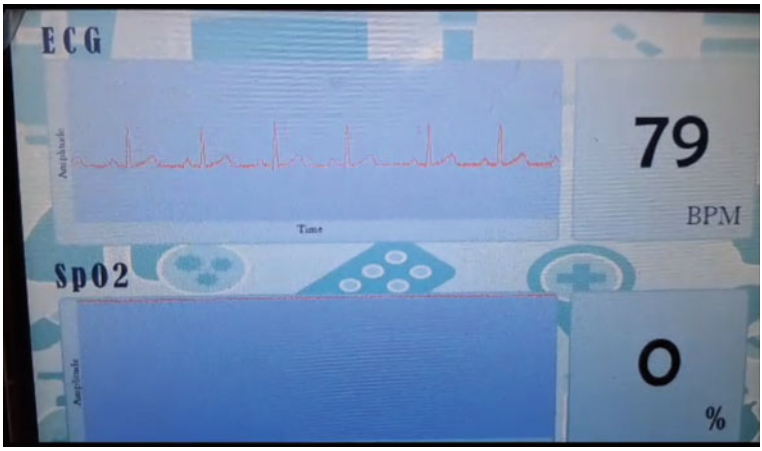


Fig. 5 The display on the module uses a LCD TFT when monitoring

In the web view, the top of the screen displays a graphical ECG signal that clearly depicts electrical heart activity. In addition, users can also see real-time BPM (heart rate per minute) values, which provide important information about the state of the heart. At the bottom of the display, there is a graphic of the SpO2 signal and the SpO2 value which measures the oxygen level in the blood. It provides comprehensive monitoring of two important parameters in health monitoring.

Apart from real-time monitoring, this module also allows users to save monitoring data easily and automatically in excel format. This is very useful in recording data as well as in long-term analysis. With this capability, users can track their health progress over time and provide valuable information for medical or research purposes. Thus, this web interface becomes a powerful tool for monitoring and analyzing comprehensive health data.

Testing of this tool is carried out to determine whether the value read on the tool is accurate with reference values from the calibrated Pulse Oximeter. Here are the results of data retrieval 10 times in modules and comparison tools:

From the results of Table 1 above, you can see the results of data collection 10 times in modules and comparison tools. The average reading of SpO_2 values in the module is 96.2 while the reading of SpO_2 values in the comparison tool is 96.9. Of the 10 times data retrieval, the largest error value was 0.02% when the module showed a value of 96 and in the comparison showed a value of 98. While the smallest error value is 0% when the SpO_2 value shown in the module and comparison tool is the same. From these results, it can be said that the module has a good accuracy of reading SpO_2 values.

After testing related to the accuracy of reading values by comparing the research module with a comparison tool, namely Pulse Oximeter, then testing of respondents can be done to obtain data before being sent and after being sent on the HTML web page. The data will later be used to analyze data loss and time delay from sending

Table 1 Results of comparison of SpO_2 values in modules and comparison tools

Data retrieval To-	SpO_2 value on module (%)	SpO_2 value on comparison tool (%)	Error (%)
1	96	97	0,010,309
2	96	97	0,010,309
3	97	98	0,010,204
4	96	97	0,010,309
5	97	98	0,010,204
6	96	96	0
7	96	96	0
8	96	97	0,010,309
9	96	98	0,020,408
10	96	95	-0,01,053
Average	**96,2**	**96,9**	**0,007,153**

data on HTML web pages. The data collection process was carried out on 10 male respondents with an age range of 20–30 years for 15 s using the MAX30100 module which will then be processed and sent by the Wemos Mega 2560 microcontroller. Data sent by the module can be viewed via the TFT LCD and will be recorded through the Delphi 7 application, while data received by the HTML Web Page will be viewed and recorded from the Web.

Latency tests are performed to determine the latency of the delivery system between the module and the site. The delay will indicate how long it takes to send data from the sender to the receiver in seconds or microseconds. This time may be affected by distance, physical media, network, or long processing time. In sending SpO_2 signals, this time delay is important because it ensures that the system can monitor the signal in real-time. This is because there are several conditions that require the state of the SpO_2 signal to be known in real-time so that the best possible handling can be done. In Table 2. Here are the standard categories for time delay [27, 28]:

It can be seen in Table 3, the highest time delay value was found in the test in the seventh respondent, which was 307 ms. As for the smallest time delay found in the test in the ninth respondent, which was 154 ms. For the average value of time

Table 2 Standard category of time delay. This standard category will later become a reference for determining performance results in the module delivery process based on the time delay value

Time Delay Category	Target
Very good	<150 ms
Good	150–300 ms
Medium	300–450 ms
Bad	>450 ms

Table 3 The results of the time delay value on the module

No	Respondents	Time Delay Value (ms)
1	1	195
2	2	214
3	3	236
4	4	235
5	5	207
6	6	301
7	7	307
8	8	167
9	9	154
10	10	177
Average		**219,3**

delay in the module, it was 219.3 ms. Hereby it is said that the delivery system in the module is very good at sending SpO_2 signals.

With the value of data loss and time delay, there are differences in the amount of data sent and received, as well as time differences in the data. Therefore, if the SpO2 signal graph of those sent and received is compared to Fig. 6, it will show the difference in the amount of data and the time difference between the two.

In Table 4 is the data sent and received data obtained from the fourth respondent. Based on the results of data analysis using the t-test, there is no significant difference between the data sent and the data received where the obtained p-value = 0.835604532 and the value of the specified significance level of 0.05, it can be seen that the p-value > 0.05. Therefore, hypothesis H0 is accepted, which states that there is no significant difference between the data sent and the data received and it can be said that the sending system in the module is very good at sending SpO2 signals on the HTML Web Page. Analysis is conducted to assess the module's manufacturing

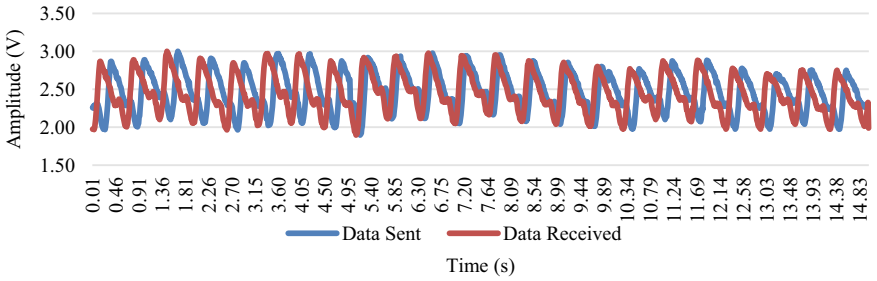


Fig. 6 Comparison graph of SpO_2 signal data sent and received to one respondent. It can be seen that there is a time difference between the data sent and received, this is due to the time delay of the sending process

Table 4 T Test analysis of data sent and data received

t-Test: Two-Sample Assuming Equal Variances		
	DATA SEND	DATA RECEIVED
Mean	2,456,454,871	2,458,416,422
Variance	0,066,946,626	0,067,051,602
Observations	1500	1500
Pooled Variance	0,066,999,114	
Hypothesized Mean Difference	0	
df	2998	
t Stat	-0,207,537,081	
P(T < = t) one-tail	0,417,802,266	
t Critical one-tail	1,645,362,047	
P(T < = t) two-tail	0,835,604,532	
t Critical two-tail	1,960,755,583	

performance. The purpose of this study was to evaluate the transmission quality of the SpO2 signal data on the HTML web page.

4 Discussion

After testing the module, data collection and analysis are performed to determine the module's production performance. This study also aimed to analyze the data transmission quality of SpO2 signal on HTML web page. After conducting a test study sending signal data and SpO2 values of 10 respondents to an HTML web page, the highest time delay value was found in the test on the seventh respondent, which was 307 ms. While for the smallest time delay there was a test on the ninth respondent, which was 154 ms. For the average time delay value in the module, it was 219.3 ms. Delay will indicate how long it takes to send data from sender to receiver in seconds or microseconds. This time may be affected by distance, physical media, network, or long processing time. Hereby it is said that the delivery system on the module is very good at sending SpO2 signals.

This study was compared with the research of Lanny Agustine et al., who conducted research on the development of arrhythmia monitoring tools using pulse oximeters through the Blynk application. In this study, using an IoT delivery system using Wi-Fi to send data to the Blynk application. However, this study still does not display signal data [17]. The result of this research is that sending signals can be done on HTML Web Pages. So, this research has advantages in monitoring not only BPM values and SpO2 values, but also ECG signals and SpO2 signals.

In making modules in this study, of course, there are still limitations which can later be used as references for development. This module can only be accessed on the

same Wi-Fi so the connection distance is of course limited and cannot be accessed too far. The sensor user also uses the MAX30100 with a reflectance sensor type, so unlike the transmittance circuit, sometimes the MAX30100 sensor is not compatible with all types of SpO2 calibrators.

The results of this study are that it is possible to develop a SpO2 monitoring system through an HTML Web Page with display values and signals that can be accessed using Wi-Fi. With communication via Wi-Fi, monitoring can be done remotely as long as it's on the same Wi-Fi. This module is also easy to use because it does not require cable installation or special applications, only using a browser on each personal device. That way, monitoring of patients can be carried out periodically even though there is no need for face-to-face meetings. This research can be a solution to a patient's vital sign monitoring system, especially SpO_2 which can be carried out periodically in sterile conditions because it can be done anywhere as long as you have the same Wi-Fi access.

5 Conclusion

This study aims to analyze the delivery of the module to the SpO2 value and signal data on HTML Web Pages. The delivery analysis is assessed based on the value of data loss and time delay during the SpO2 data transmission process. Based on the obtained data, the SpO2 transmission signal can be considered as good. Indeed, if you look at the results of the data analysis, the average of the time delay value on the module is 219.3 s.

It is hoped that this research can be extended by adding a wider range of network connections available online so that it can be accessed over longer distances. The sensor module used is of the reflective type, so in the future one can expect to use a sensor with a transmission type.

References

1. Dias D, Cunha JPS (2018) Wearable health devices—vital sign monitoring, systems and technologies. Sensors (Switzerland) 18(8). https://doi.org/10.3390/s18082414
2. Firdaus MA, Pudji A, Mak MR (2020) Design a vital sign monitor for body temperature (Axilla) and Oxymetry parameters. J Electron Electromed Med Inf 2(2):58–64
3. Kadhim KT, Alsahlany AM, Wadi SM, Kadhum HT (2020) Monitor human vital signs based on IoT technolgy using MQTT protocol. AIP Conf Proc 2290. https://doi.org/10.1063/5.0027363
4. Khan Y, Ostfeld AE, Lochner CM, Pierre A, Arias AC (2016) Monitoring of vital signs with flexible and wearable medical devices. Adv Mater 28(22):4373–4395. https://doi.org/10.1002/adma.201504366
5. Sameh R, Genedy M, Abdeldayem A, Abdel Azeem MH (2010) Design and implementation of an SPO2 based sensor for heart monitoring using an android application. J Phys Conf

Ser 1447(1). https://doi.org/10.1088/1742-6596/1447/1/012004. Author F (2010) Contribution title. In: 9th International proceedings on proceedings, pp 1–2. Publisher, Location
6. Suherlim DA, Permana H, Lubis L (2018) Correlation between haemoglobin concentration and oxygen saturation (SpO2) in elderly professors. J Thee Med Sci (Berkala Ilmu Kedokteran) 50(2):157–162. https://doi.org/10.19106/jmedsci005002201804
7. Vold ML, Aasebø U, Wilsgaard T, Melbye H (2015) Low oxygen saturation and mortality in an adult cohort: the Tromsø study. BMC Pulm Med 15(1):1–12. https://doi.org/10.1186/s12890-015-0003-5
8. Shepherd SJ, Pearse RM (2009) Role of central and mixed venous oxygen saturation measurement in perioperative care. Anesthesiology 111(3):649–656. https://doi.org/10.1097/ALN.0b013e3181af59aa
9. Shah S et al (2020) Novel use of home pulse oximetry monitoring in COVID-19 patients discharged from the emergency department identifies need for hospitalization. Acad Emerg Med 27(8):681–692. https://doi.org/10.1111/acem.14053
10. Najwa U, Kamaludin I, Ramli NI (2022) IoT patient monitoring system for COVID-19. Evol Electr Electron Eng 3(1):598–602
11. Su J et al (2010) Real-time fusion of ECG and SpO2 signals to reduce false alarms. Comput Cardiol, 1–4. https://doi.org/10.22489/CinC.2018.163
12. Deshmane AV (2009) False arrhythmia alarm suppression using ECG, ABP, and photoplethysmogram. Mit, 1–93. http://dspace.mit.edu/handle/1721.1/54209
13. Sola A et al (2014) Safe oxygen saturation targeting and monitoring in preterm infants: can we avoid hypoxia and hyperoxia? Acta Paediatr Int J Paediatr 103(10):1009–1018. https://doi.org/10.1111/apa.12692
14. Shenoy N, Luchtel R, Gulani P (2020) Considerations for target oxygen saturation in COVID-19 patients: are we under-shooting? BMC Med 18(1):1–6. https://doi.org/10.1186/s12916-020-01735-2
15. Tripathi A, Tyagi VK, Vivekanand V, Bose P, Suthar S (2020) Challenges, opportunities and progress in solid waste management during COVID-19 pandemic. Case Stud Chem Environ Eng 2:100060. https://doi.org/10.1016/j.cscee.2020.100060
16. Toombs CS, Boody BS, Bronson WH, Girasole GJ, Russo GS (2021) Safe spine surgery during the COVID-19 Pandemic. Clin Spine Surg 34(3):87–91. https://doi.org/10.1097/BSD.0000000000001084
17. Kadarina TM, Priambodo R (2021) Performance evaluation of IoT-based SpO2 monitoring systems for COVID-19 patients. J Electron Electromed Eng Med Inf 3(2):64–71. https://doi.org/10.35882/jeeemi.v3i2.1
18. Agustine L, Muljono I, Angka PR, Gunadhi A, Lestariningsih D, Weliamto WA (2018) Heart rate monitoring device for arrhythmia using pulse oximeter sensor based on android. 2018 international conference computer engineering network intelligent. multimedia, CENIM 2018—Proceeding, pp 106–111. https://doi.org/10.1109/CENIM.2018.8711120
19. Luthfiyah S, Ramadhani ER, Indrato TB, Wongjan A, Lawal KO (2022) Vital signs monitoring device with BPM and SpO2 notification using telegram application based on Thinger.io platform. Indones J Electron Electromed Eng Med Inf 4(1):1–7. https://doi.org/10.35882/ijeeemi.v4i1.1
20. Sidhartha G (2021) IoT based pulse oximeter. Int J Res Appl Sci Eng Technol 9(8):2946–2951. https://doi.org/10.22214/ijraset.2021.37782
21. Gade P (2021) IoT Based pulse oximeter using ESP8266. SSRN Electron J. https://doi.org/10.2139/ssrn.3918115
22. Tham OY, Markom MA, Bakar AHA, Tan ESMM, Markom AM (2020) IoT health monitoring device of oxygen saturation (SpO2) and heart rate level. Proceeding—1st international conference information technology advance mechanical electrical engineering. ICITAMEE 2020, pp 128–133. https://doi.org/10.1109/ICITAMEE50454.2020.9398455
23. Ganesh KVSS, Jeyanth SPS, Bevi AR (2022) IOT based portable heart rate and SpO2 pulse oximeter. HardwareX 11:e00309. https://doi.org/10.1016/j.ohx.2022.e00309

24. Onubeze A (2016) Developing a wireless heart rate monitor with MAX30100 and nRF51822. Helsinki Metrop Univ Appl Sci., 13. http://urn.fi/URN:NBN:fi:amk-2016120118617
25. Fadjar Bastari W, Arif Anshori M, Ramadhani MR (2022) Design & construction of oxygen saturation measuring equipment through the thingspeak server with max 30100 sensor. BEST J Appl Electr Sci Technol 4(2):75–80. https://doi.org/10.36456/best.vol4.no2.6155
26. Shamsuddin S, Argenan K, Kamat SR, Nordin MH (2018) Development of a simple mobile robot for humanrobot interaction in health care environment. J Adv Manuf Technol 12(Specialissue4):159–172
27. Sugeng W, Istiyanto JE, Mustofa K, Ashari A (2015) The impact of QoS changes towards network performance. Int J Comput Netw Commun Secur 3(2):48–53. http://www.ijcncs.org/published/volume3/issue2/p5_3-2.pdf
28. Chen Y, Bindel D, Song H, Katz RH (2004) QoS Requirements of network applications on the internet. Comput Commun Rev 34(4):55–66. https://doi.org/10.1145/1030194.1015475

Wide Communication Coverage ECG—Lead II Monitoring Using Local Host HTML Web Page



Priyambada Cahya Nugraha, Nurdiansyah Wahyu Bima Putra, I. Dewa Gede Hari Wisana, Moch Prastawa Assalim Tetra Putra, Riqqah Dewiningrum, and Divanda Natya Kirana

Abstract Cardiovascular disease one of the leading global causes of death, this need to be monitored for heart conditions which must done regularly. Electrocardiograph (ECG) method currently being developed facilitate patient monitoring, and it falls under the Internet of Things as we know IoT advancements for ECG devices. This study for analyze ECG signal and BPM also called Beats per Minutes values sent and received through IoT media to aid in the diagnostic process. The study's contribution lies in understanding the format of ECG signals and BPM values transmitted and received via IoT media. The AD8232 sensor is used in the process, and its output is processed and displayed on an HTML web page. The research shows an average data loss of 0.3652%. Based on this study, the transmission of ECG signal data and BPM values on HTML web pages is deemed effective. The findings of this study are expected to be utilized on further research and may be beneficial for developing an ECG monitoring system to improve patient care, especially for those requiring sterile treatment.

Keywords AD8232 · ECG · HTML · Data loss

1 Introduction

One of the characteristics of the vital sign monitor is the electrocardiogram, or ECG. Other measures, besides the ECG, include blood pressure, SpO2, respiration rate, and temperature [1]. ECG and SpO2 can be used for monitoring and diagnostics to find the existence of a disease [2]. The normal ECG signal consists of PQRST waves [3]. The morphology of the heart's electrical waves and the amplitude values of the

P. C. Nugraha · N. W. B. Putra · I. D. G. H. Wisana (✉) · M. P. A. T. Putra · R. Dewiningrum · D. N. Kirana
Department of Electromedical Engineering, Health Polytechnic Ministry of Health Surabaya, Surabaya, Indonesia
e-mail: hariwisana@yahoo.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_17

P, QRS and T waves are the main results in examining the condition of the heart [4]. There are two types of information provided when recording an ECG, namely measuring the time interval on the ECG so that it can show whether the electrical activity that occurs in the heart is normal, slow, fast, or irregular. Apart from that, ECG recording is also done to measure the amount of electrical activity passing through the heart muscle, so that cardiologists can find out whether there are parts of the heart that are too big or working too much [5]. Cardiovascular disease one of top causes mortality globally [6]. This prompts the development of electrocardiograph device research [7]. 17.7 million individuals worldwide passed away because heart disease at 2017, making up 31% deaths, according data analysis from World Health Organization. Heart disease alone causes 37% of all deaths in Indonesia [8]. Periodic cardiac function monitoring is necessary to prevent patient harm since, in the event of neglect, delayed treatment may result in patient harm [9]. If there is interference with cardiac action, the importance of heart function can be fatal [10, 11]. Special management was required, particularly during the COVID-19 pandemic, to prevent the disease from spreading further [12, 13] and to prevent the examination of patients with cardiac disease from being disrupted. IoT (Internet of things) can be used in efforts to better handle and provide cardiac patients [14, 15]. This IoT system makes it easy for users and other system to connect and communicate with each other, it can have a solution for this problem [16, 17].

On 2016, Zhe Yang et al. did research from Internet of Things-based Wearable ECG monitoring system this displays ECG signal data via the HTTP and MQT protocols. This study discovered that ECG signals transmitted over Wi-Fi could be seen at web server that was accessible from any device. But the outcomes still not entirely correct [18]. In 2019 saw the completion of research in IoT-based ECG monitoring to Smart Healthcare by Tamanna Shaown et al. In this study, ECG signals are captured use the AD8232 module and delivered through Wi-Fi from Raspberry Pi, which serves a microcontroller as signal processing. However, the application for showing these signals is not described in this paper, and no ECG signal data [19]. In 2020, Alamsyah et al. conduct research at vital sign monitoring with parameters of temperature, heart rate and blood pressure. This study makes use of a Raspberry Pi and wifi to show data that will be stored in a database online. Even if just the numbers and graphs of these values are provided, the accuracy of the findings on the web is extremely exact. Additionally, SpO2 and ECG values are not utilised [20]. On 2021, S. Hadiyoso et al. conducted research on vital sign monitoring applications with SpO2, ECG, BPM, and temperature parameters. In this study, Bluetooth is used as a communication medium with mobile phones and the cloud as a database system. However, it is not explained how the transmission of data from the device to the application used and the use of the application requires storage on the device for application installation [21]. Agatha P. J. S. et al. used the Internet of Things to do research on vital signs with temperature and respiration data in 2021. The platform employed in this study is thinger.io, which will notify Telegram if specific conditions are met. The sensors utilized are the piezoelectric respiration sensor and the DS18B20 temperature sensor. However, the web-based display is currently merely a value without a signal [22]. In 2022, Sari L. et al., conduct research on monitoring BPM

and SpO2 sent via wifi which will then be displayed on the tinger.io platform and will display notifications on telegrams when certain conditions occur. The sensor used in this study uses MAX30100 to read BPM and SpO2. However, this study only displays values on web and telegram displays, no displaying signals [23]. 2019 saw the completion research at ECG monitoring utilizing the IoT system shown on Thingspeak platform by Pagalada P. K. et al. In this investigation, graphs of the patients' heart rates and BPM values were shown. However, neither the ECG signal's shape nor an analysis of the obtained BPM value's error value are displayed or provided [24]. In 2022, Dewi Rushalina et. al., conducting research on sending and receiving ECG data through an IoT system using wi-fi. This study uses the AD8232 sensor to detect 3 Lead ECG with the ESP32 microcontroller. The sent signal can be displayed on a computer or cell phone via HTML media. However, the wi-fi network used must use the same network and there is no display on the device [25].

Therefore, it can be inferred from the many research mentioned above that not much is actually displayed by the ECG monitoring device. In fact, as is well known, the ECG signal and value include significant data on the condition of the human body. These studies primarily made use of already-existing websites and supplied data exclusively. In order to display the value and signal of the ECG in real time, we will attempt to develop our own website for tracking ECG data using HTML Web Pages. The AD8232 sensor will also be employed in this investigation since, according to some of the papers mentioned above, it can be used as a sensor module to record ECG in humans.

Based on these issues, the authors plan to study ECG Lead II Monitoring Via HTML Web Page using ECG Lead II parameters that will be shown on an IoT platform, namely an HTML web page. This study examines the transmission of ECG signals and BPM readings both before and after the transmission of IoT. With the help of research, believed that ECG monitoring system signal display may improved so physicians or nurses provide better patient care.

2 Materials and Method

Phantom ECG was employed as a module comparison tool in this study, which was carried out at Electronics Laboratory of the Department of Electromedical Technology, Poltekkes Kemenkes Surabaya. If it matches the comparator, then module testing will be carried out on 10 male respondents with an age range of 20–30 years. Data collection will be carried out for 15 s for respondents who are calm and in a sitting position. The data sent is recorded via Delphi 7, while the data receive via HTML Web Page is then compared to look for data loss analysis.

This study, the tool was evaluated using Phantom ECG to see whether Phantom ECG and the module worked well together once the tool design was complete. 10 respondents are used in the data gathering process if the findings are suitable. The module will keep track of the values and signals provided by the responders, which are subsequently delivered to an HTML web page and displayed on a TFT LCD.

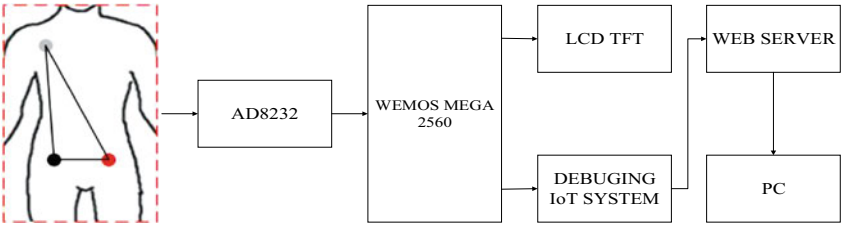


Fig. 1 Block diagram of the system from ECG–lead II monitoring module via HTML

The Delphi 7 application will be used to retrieve data from the module and save it in excel format. The collecting of data from HTML Web Pages will be done online and stored in an excel format. The data received by the HTML Web Page and the data supplied by the module will then be compared, and the data loss will be examined.

The ECG data that was collected prior to transmitting is the research's independent variable. The ECG data that has been sent is the dependent variable. In this study, the Wemos Mega 2560 microcontroller is the controlled variable.

The AD8232 sensor was used investigation for collect ECG data from respondents. Wemos Mega 2560 microcontroller is utilized to process the data. The Nextion 5 Inch TFT LCD Enhanced Models are used for the display [26].

This paragraph will explain Fig. 1 that is block diagram of system. A system block diagram consisting of three main components namely input, process and output. The ECG data collected before transmission is the independent variable of the study. The ECG data that has been sent is the dependent variable. In this research, the Wemos Mega 2560 microcontroller is the controlled variable.

In this research, the input used is the AD8232 sensor as a means of capturing ECG Lead II data from the respondent's body. The Wemos Mega 2560 microcontroller is used to process data, where data from the sensor output will be processed so that ECG Lead II signal data and BPM values can be sent. The Wemos Mega 2560 microcontroller will send ECG signal data and BPM values and will then be shown on the Nextion 5 Inch TFT LCD as a display on the device and then to the IoT platform that has been designed on a PC and will be stored on the PC.

Whereas in Fig. 2 depicts a flowchart that explains how this module system works in detail. When the user wants to activate the module, the first step is to press the ON button to turn it on. After the ON button is pressed, the module will start the initialization process and make preparations to carry out monitoring procedures. After the power-on procedure is complete, the next step is to receive input from the AD8232 sensor, followed by detection of the ECG signal and BPM value. Data in the form of ECG signals and BPM values will be displayed on the TFT LCD.

This system also has a feature that allows adaptation to Wi-Fi network conditions where if the device is not connected to Wi-Fi, the ECG signal and BPM value will still be displayed on the TFT LCD screen. However, if the device is connected to Wi-Fi, the Wemos Mega microcontroller will send ECG signals and BPM values

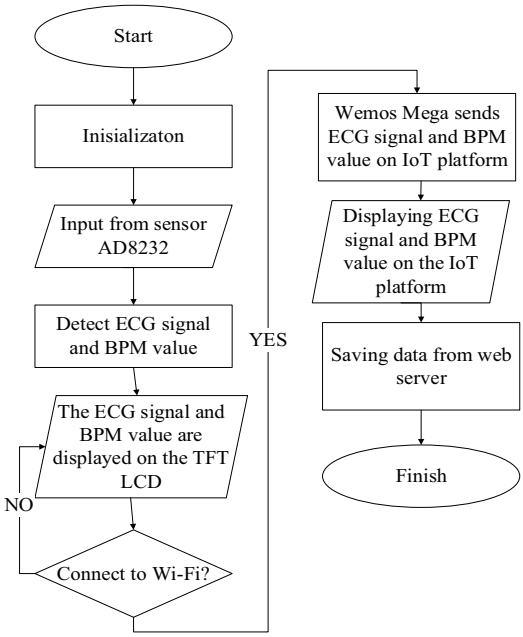


Fig. 2 The system flowchart from ECG–lead II monitoring module via HTML web page

to the IoT platform for display allowing users to access the data via the web. The system will be complete.

2.1 Data Analysis

This formula will be applied while getting BPM value data, which will be repeated six times to determine the average BPM value Eq. (1) contains the average value's formula:

$$\overline{x} = \frac{x1 + x2 \cdots + xn}{n} \tag{1}$$

A system error is indicated by %error. The average difference between each data point is the lower Error value. Errors might be a symptom that the design or model deviates from the standard Eq. (2) displays the error formula.

$$\%\text{error} = \frac{(\boldsymbol{x_n} - \mathbf{x})}{\boldsymbol{x_n}} \times 100\% \tag{2}$$

xn is the measured value of the motor calibrator. X for design measurement value.

3 Result

Module testing was carried out with a comparison in the form of Phantom ECG and testing of 10 respondents. Figure 3 is the result of the design of this research module:

In Fig. 3 is the design of this research module. This module uses a power bank as the main power supply source for the entire network, enabling potential and convenience in use. The AD8232 sensor series functions to collect ECG data from the body which will then be forwarded to the microprocessor. The Wemos Mega 2560 microprocessor will process the sensor output. Apart from processing the output from the AD8232 sensor, the Wemos Mega 2560 is used to input web designs, transfer data to the web, and save data from the web.

Then the processing results from the microprocessor will also be displayed on the LCD screen display and HTML web page. The Nextion 5 Inch TFT LCD is the display used in the module. The screen presents ECG and BPM data clearly and informatively, providing very useful information in understanding the patient's heart condition.

In Fig. 4 shows the web interface used in the monitoring process. With internet access, users can easily access the ECG and BPM signal display from this module, either via smartphone or computer. To begin monitoring, users need to enter the appropriate IP address on their web browser, the IP address entered must match the IP address that appears on the Nextion 5 Inch TFT LCD display. It allows users to monitor heart conditions and measure their heart easily and accurately.

In the web view, the top of the screen displays a graphical ECG signal that clearly depicts electrical heart activity. In addition, users can also see real-time BPM (heart rate per minute) values, which provide important information about the state of the heart. At the bottom of the display, there is a graphic of the SpO2 signal and the SpO2 value which measures the oxygen level in the blood. It provides comprehensive monitoring of two important parameters in health monitoring.

Apart from real-time monitoring, this module also allows users to save monitoring data easily and automatically in excel format. This is very useful in recording data as well as in long-term analysis. With this capability, users can track their

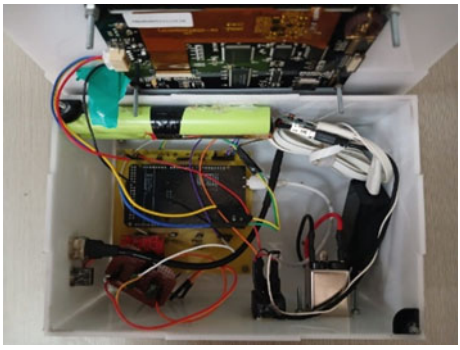


Fig. 3 Circuit results on ECG–lead II monitoring module

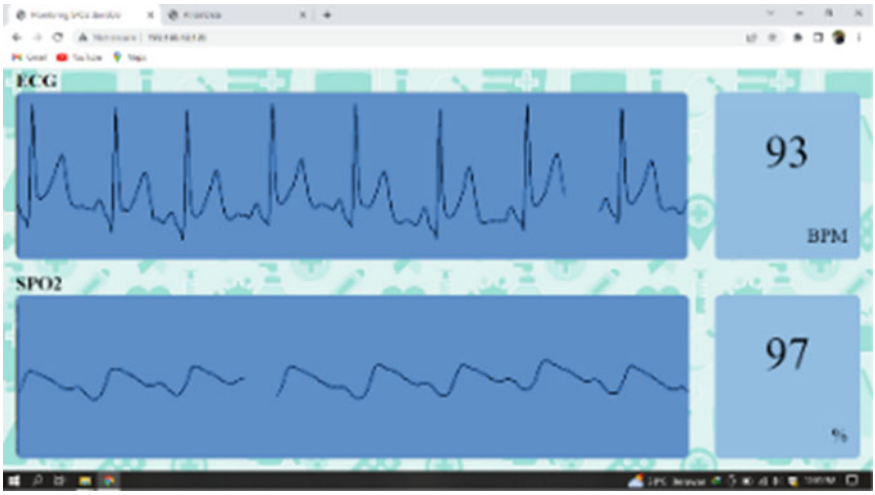


Fig. 4 Web display when monitoring

health progress over time and provide valuable information for medical or research purposes. Thus, this web interface becomes a powerful tool for monitoring and analyzing comprehensive health data.

In Fig. 5 is web display when monitoring. In the initial display, the IP Addres will appear, then when the LCD screen is pressed, a display will appear as shown above. Monitoring data will appear automatically when the device is attached to the patient. The data display is in the form of ECG signals and BPM values.

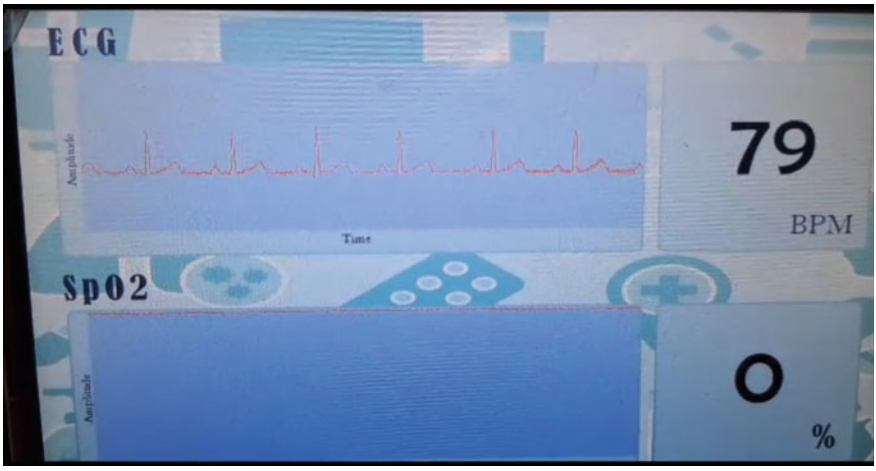


Fig. 5 TFT LCD when monitoring

At the top there is a graphical display of the ECG signal and BPM value. Then at the bottom there is a graphical display of the SpO2 signal and SpO2 value. Then the monitoring data can be saved automatically to Excel.

Testing of this tool is carried out to find out whether the value read on the tool is accurate with the reference value from the Phantom ECG. This test is carried out by recording the BPM value that appears on the module compared to the BPM value that appears on the Phantom ECG 6 times. The BPM used from the ECG Simulator is 60 BPM, 80 BPM, 100 BPM and 120 BPM. The results of the data collection will later be analyzed using an error value to view the module's BPM value. The following is the result of collecting data 6 times at module and appeal equipment:

From the results of Table 1 above can see results of collecting data 6 times on the module and Phantom ECG. The average result of reading value of 60 BPM on the module is 59.67 so the resulting error value 0.0056%. Reading 80 BPM value on the module 79.17 so the resulting error value 0.0104%. Reading of the value of 100 BPM on the module 99.83 so the resulting error value 0.0017%. Reading of the 120 BPM value on the module 119.67 so the resulting error value 0.0028%. Of the 6 times data collection, the largest error value 0.0104% when set Phantom ECG at a BPM value of 80. While the smallest error value 0.0017% when setting the ECG Simulator at a BPM value of 100. From these results, it can be said that the module has good BPM value reading accuracy.

After checking the accuracy of the value reading by comparing the search module with the comparator, namely Phantom ECG, then testing of respondents can be done to obtain data before it is sent and after it is sent on the HTML web page. This data will later be used to analyze data damage from sending data on HTML web pages. Data collection action was undertaken on 10 male responden range age 20–30 years for 15 s with the AD8232 module which will then be processed and sent by the Wemos Mega 2560 microcontroller. Data sent by the module can be viewed via the TFT LCD and will be recorded. through the Delphi 7 application, while the data received by the HTML Web Page will be viewed and recorded from the Web.

Data loss tests seek to ascertain the amount of data lost during the delivery system's transition from the module to the HTML Web Page. How much data is lost during data transfer is shown by this data loss. This data loss indicates a problem with data packet transmission, which can be brought on by a number of things, including network overload, problems with the physical medium, problems with the receiving end, etc.

Table 1 Results of Comparison of BPM Values in Modules and Comparison Tools

Setting	BPM value on module						Average	Error (%)
	1	2	3	4	5	6		
60 BPM	59	60	60	60	59	60	**59.67**	**0.0056**
80 BPM	79	80	79	79	79	79	**79.17**	**0.0104**
100 BPM	100	100	99	100	100	100	**99.83**	**0.0017**
120 BPM	120	120	120	120	119	119	**119.67**	**0.0028**

Table 2 Standard category of data loss

Data loss category	Target (%)
Incredible	0
Nice	3
Currently	15
Poor	25

Data loss in signal data is significant because it will change how the signal is perceived on the medium. The amount of lost data increases with the value of the data loss. This alters the signal's physical shape from what it should be and might lead us to lose crucial information in it. In Table 2 below are the standard categories of data loss [27, 28]:

In Table 2 the standard category of data loss. The standard category of data loss will later become a reference to determine the performance results in the module delivery process. Testing the value of data loss is done by finding the percentage of the difference between the data sent and the data received for 15 s. Data loss is divided into four categories, namely very incredible, nice, currently, and poor.

When the lost data value is at 0%, it is included in the very incredible category and if the data loss value is at 3%, it is included in the good category. Meanwhile, when lost data is at 15%, it is included in the currently category. Then if the lost data is at 0%, it is included in the poor category.

Based on Table 3 above is the value of data loss from 10 respondents. The amount of data sent has a different amount of data from each respondent with a range of 1503 to 1508 data but the amount of data received has the same amount in each respondent, namely 1500 data. The high data loss value is found in the fourth respondent test, which is 0.5305%. While the smallest data loss is found in the ninth and tenth respondent tests, which amounted to 0.1996%. The average loss value in this module is 0.3652%. So it can be seen that the transmission system in the module is very good at sending ECG signals.

In Fig. 6 there is a graph that is the result of a comparison plot of ECG Lead II signal data sent and received on the HTML Web Page for 15 s. It can be seen that there is a difference in the amount of data sent from the module and received by the HTML Web Page. The data sent from the module amounted to 1505 data, while the data received amounted to 1500 data.

The difference in the amount of data indicates data loss during the transmission process. Then, there is also a time difference in the data sent and received data. This is due to the time delay in the transmission process. In reading the BPM value between module and web readings is the same, namely 90 BPM.

In Table 4 is the data sent and received data obtained from the fourth respondent. Based on the results of data analysis using the t test, there is no significant difference between the data sent and the data received where the obtained p-value = 0.50461 and the value of the specified significance level of 0.05, it can be seen that the p-value > 0.05. Therefore, the hypothesis H0 is accepted, which states that there is no significant difference between the data sent and the data received and it can be

Table 3 The results of the data loss value on the module

No	Respondents	Amount of data sent	Amount of data received	Data loss value (%)
1	1	1506	1500	0.3984
2	2	1505	1500	0.3322
3	3	1504	1500	0.2660
4	4	1508	1500	0.5305
5	5	1505	1500	0.3322
6	6	1507	1500	0.4645
7	7	1507	1500	0.4645
8	8	1507	1500	0.4645
9	9	1503	1500	0.1996
10	10	1503	1500	0.1996
Average				**0.3652**

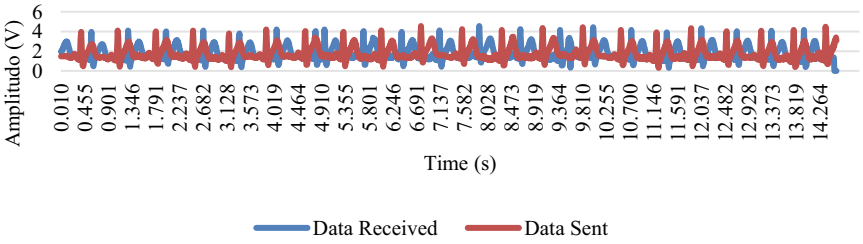


Fig. 6 ECG signal data provided and received for one respondent are shown in a comparison graph. The different is the amount of data received, which indicates data loss

said that the sending system in the module is very good at sending ECG signals on the HTML Web Page. analysis are conducted to assess the module's manufacturing performance. The purpose of this study was to evaluate the transmission quality of the ECG signal data on the HTML Web Page.

4 Discussion

Once the module has undergone testing, data collection and result analysis are conducted to assess the module's manufacturing performance. The purpose of this study was to evaluate the transmission quality of the ECG signal data on the HTML web page. The research tests involved sending ECG signal data and BPM values from 10 respondents on HTML web pages for a 15-s test duration. There is a difference in the amount of data sent and received based on search results. While the BPM values remain the same between sent and received data, there is a time lag during

Table 4 T Test Analysis of Data Sent and Data Received (alpha = 0.05)

T Test: Two Independent Samples									
Summary					Hyp mean diff 0				
Groups	Count	Mean	Variance	Cohen d					
Data Sent	1508	1.7259	0.1682						
Data Received	1500	1.7160	0.1615						
Pooled			0.16492	0.024335					
T Test: Equal varianess				Alpha	0.05				
	std err	t-stat	df	p-value	t-crit	Lower	Upper	Sig	Effect R
One Tail	0.0148	0.6673	3006	0.2523	1.6453			no	0.0121
Two Tail	0.0148	0.6673	3006	0.5046	1.9607	-0.0191	0.0389	no	0.0121

transmission, resulting in a time difference between the module and the HTML Web Page. The data loss testing during the shipping process showed varying results for different respondents, with the highest data loss at 0.5305% for the fourth respondent and the smallest at 0.1996% for the ninth and tenth respondents. Average data loss value for the module is 0.3652%, indicating the send system at module efficient on transmitting ECG signal data.

This research is compared to the study by Pagalada P. K. et al., based on ECG monitoring with IoT system available on Thingspeak. In previous studies, the heart rate graph and BPM value were displayed in the patient, but it did not show the shape of the ECG signal and analyze the error value of the received BPM value [21]. This study shows this research has advantages in monitoring not only BPM value, but also ECG signals. So that, the sending ECG signals can be done successfully on HTML web pages.

However, the research module does have some limitations. One limitation is the Wi-Fi access, which is only accessible on the same Wi-Fi network, limiting the connection distance. The sensor used, AD8232, can only be used on one lead, resulting in significant noise and instability in the signal's baseline when used on a human patient.

The system developed in this module aims to analyze ECG signals and BPM values before and after sending them on HTML web pages. The hope is that this research can lead to the development of an ECG monitoring system with display values and signals, allows regular ECG monitoring without direct patient monitoring and can be used anywhere with a Wi-Fi network. Module's simplicity, as it only requires a browser on individual devices, without expensive cable setups or specialized software, makes it user-friendly for periodic patient checks without the need for in-person sessions. This research provides a potential solution for a patient's vital sign monitoring system, especially at sterile settings where ECG can be periodically performed.

5 Conclusion

The objective this study was investigate transmission of BPM values and ECG signals from the module to HTML web pages. The analysis of the data revealed that the quality of ECG signal delivery can be considered good, based on the obtained data.

The data analysis showed that the average data loss in the module is 0.3652%. This indicates sending system in module highly effective in transmitting ECG signal data, as both the Data Loss values fall under the "Very Good" category in the Standard Category.

The findings of this research are expected to be beneficial for health workers in implementing an ECG monitoring system that can be conducted periodically without the need for continuous physical contact with patients. This aspect is especially useful when patient handling requires a sterile environment. Further development of the research could involve adding network connectivity to expand the access range and incorporating up to 12 ECG leads.

References

1. Aditya Mahendra Oka G, Pudji A (2021) Design of vital sign monitor with ECG, BPM, and respiration rate parameters. Indones J Electron Electromed Eng Med Inf 3(1):34–38. https://doi.org/10.35882/ijeeemi.v3i1.6
2. Su J et al. (2010) Real-time fusion of ECG and SpO2 signals to reduce false alarms. Comput Cardiol, 1–4, 2018. https://doi.org/10.22489/CinC.2018.163
3. YN, Fitri Nurul Auliaand Fu'adah MR (2023) Fetal ECG signal processing using one-dimensional convolutional neural network (1D CNN) for fetal arrhythmias detection. In: Proceeding of the 3rd international conference on electronics, biomedical engineering, and health informatics, pp 75–87
4. Setiawidayat Sabarand Rahman AY (2022) Method for obtain peak amplitude value on discrete electrocardiogram. In: Proceedings of the 2nd International conference on electronics, biomedical engineering, and health informatics, pp 97–108
5. Yoo K-S, Lee W-H, QRS complexes detection in electrocardiogram signals based on multiresolution analysis. Futur Inf Technol Appl Serv 179:153–158. https://doi.org/10.1007/978-94-007-5063-0_23
6. Mc Namara K, Alzubaidi H, Jackson JK (2019) Cardiovascular disease as a leading cause of death: how are pharmacists getting involved?. Integr Pharm Res Pract 8:1–11. https://doi.org/10.2147/iprp.s133088
7. Roth GA et al (2018) Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 392(10159):1736–1788. https://doi.org/10.1016/S0140-6736(18)32203-7
8. Komalasari R, Nurjanah, Yoche MM (2019) Quality of life of people with cardiovascular disease: a descriptive study. Asian Pacific Isl Nurs J 4(2):92–96. https://doi.org/10.31372/20190402.1045
9. Nicholls MG, Richards AM (2007) Disease monitoring of patients with chronic heart failure. Heart 93(4):519–523. https://doi.org/10.1136/hrt.2005.078519
10. Balestrino M (2021) Role of creatine in the heart: health and disease. Nutrients 13(4). https://doi.org/10.3390/nu13041215

11. Herrmann J (2020) Adverse cardiac effects of cancer therapies: cardiotoxicity and arrhythmia. Nat Rev Cardiol 17(8):474–502. https://doi.org/10.1038/s41569-020-0348-1
12. Hui DS (2017) Epidemic and emerging coronaviruses (Severe acute respiratory syndrome and middle east respiratory syndrome). Clin Chest Med 38(1):71–86. https://doi.org/10.1016/j.ccm.2016.11.007
13. Manby L et al (2022) Healthcare workers' perceptions and attitudes towards the UK's COVID-19 vaccination programme: a rapid qualitative appraisal. BMJ Open 12(2):1–8. https://doi.org/10.1136/bmjopen-2021-051775
14. Umer M, Sadiq S, Karamti H, Karamti W, Majeed R, Nappi M (2022) IoT based smart monitoring of patients' with acute heart failure. Sensors 22(7):1–18. https://doi.org/10.3390/s22072431
15. Ziryawulawo A, Ogare AC, Ayebare F, Sinde R (2022) Application of IoT and machine learning techniques for heart disease prediction and diagnosis: a comprehensive review. Int J Adv Sci Res Eng 08(07):76–85. https://doi.org/10.31695/ijasre.2022.8.7.7
16. Andalibi V, Dev J, Kim DI, Lear E, Camp LJ (2021) Making access control easy in IoT. IFIP Adv Inf Commun Technol 613:127–137. https://doi.org/10.1007/978-3-030-81111-2_11
17. Sahifa AA, Setiawan R, Yazid M (2021) Internet of things-based data delivery for remote monitoring of hemodialysis systems. J Tek ITS 9(2):A209–A214
18. Yang Z, Zhou Q, Lei L, Zheng K, Xiang W (2016) An IoT-cloud based wearable ECG monitoring system for smart healthcare. J Med Syst 40(12). https://doi.org/10.1007/s10916-016-0644-9
19. Shaown T, Hasan I, Mim MMR, Hossain MS (2019) IoT-based portable ECG monitoring system for smart healthcare. 1st International conference advance science engineering robotics technology. 2019, ICASERT 2019, vol 2019, no. Icasert, pp 1–5. https://doi.org/10.1109/ICASERT.2019.8934622
20. Alamsyah, Subito M, Ikhlayel M, Setijadi E (2020) Internet of things-based vital sign monitoring system. Int J Electr Comput Eng 10(6):5891–5898. https://doi.org/10.11591/ijece.v10i6.pp5891-5898
21. Hadiyoso S, Alfaruq A, Tulloh R, Rohmah YS, Susanto E (2021) Internet of things based real-time vital sign monitoring system using mobile application. J Appl Eng Sci 19(3):807–813. https://doi.org/10.5937/jaes0-28774
22. Santoso APJPJ, Luthfiyah S, Indrato TB, Omoogun M (2021) Vital sign monitor device equipped with a telegram notifications based on internet of thing platform. Indones J Electron Electromed Eng Med Inf 3(3):108–113. https://doi.org/10.35882/ijeeemi.v3i3.4
23. Luthfiyah S, Ramadhani ER, Indrato TB, Wongjan A, Lawal KO (2022) Vital signs monitoring device with BPM and SpO2 notification using telegram application based on Thinger.io platform. Indones J Electron Electromed Eng Med Inf 4(1):1–7. https://doi.org/10.35882/ijeeemi.v4i1.1
24. Kumar PP, Naidu TS, Vishnu S (2019) Remote Ecg monitoring system by using Iot. Int J Eng Adv Technol 9(1S5):71–73. https://doi.org/10.35940/ijeat.a1021.1291s52019.
25. Rushalina D, Wisana IDGH, Nugraha PC, Ragimova N (2022) Analysis of transmitted and received ECG signal based on internet of thing using web browser and server-client HTML protocol. J Teknokes 15(4):216–222. https://doi.org/10.35882/teknokes.v15i4.469
26. Shamsuddin S, Argenan K, Kamat SR, Nordin MH (2018) Development of a simple mobile robot for humanrobot interaction in health care environment. J Adv Manuf Technol 12(Specialissue4):159–172
27. Sugeng, Istiyanto JE, Mustofa K, Ashari A (2015) The impact of QoS changes towards network performance. Int J Comput Networks Commun Secur 3(2):48–53
28. Chen Y, Bindel D, Song H, Katz RH (2004) QoS requirements of network applications on the internet. Comput Commun Rev 34(4):55–66. https://doi.org/10.1145/1030194.1015475

Effect of Higher Temperature at Metal Plate Based on Thickness and Hardness Material Using Ultrasonic Testing Method



Ahmad Anwar Zikri Othman, Kharudin Ali, Damhuji Rifai, Nazry Abdul Rahman, Zulfikri Salleh, Muhammad Ameen Wahab, Raja Siti Nur Adiimah Raja Aris, Johnny Koh Siaw Paw, Chong Tak Yaw, Jian Ding Tan, and Talal Yusaf

Abstract Non-destructive testing (NDT) comprises various analysis techniques used in science and industry to evaluate material properties without causing harm. NDT finds applications across diverse industries, constantly evolving with new methods. It's pivotal for assessing remaining wall thickness in objects prone to corrosion or erosion, like vessel and piping. This study supporting the relativity of elevated temperature and thickness measurement in industry. While most measurements occur up to 60 °C with standard procedures, exceptions exist. Industries like refining and chemicals assess component thickness at 60 to 550 °C, where cooling isn't feasible. Temperature accelerates corrosion, demanding more frequent measurements. This study focuses on material thickness and hardness before and after heating. Stainless steel, alloy, and brass are compared, heating from 100 to 500 °C. Results reveal alloy steel thickness increases by 14.1%, followed by brass at 13.8% and stainless steel at 5.65%. In hardness, stainless steel rises by 23.8%, alloy by 23.58%, and

A. A. Z. Othman · K. Ali (✉) · D. Rifai · N. A. Rahman · Z. Salleh · M. A. Wahab · R. S. N. A. R. Aris
Faculty of Engineering Technology (Electrical and Automation), University College TATI, Telok Kalong, Jalan Panchor, 24000 Kemaman, Terengganu, Malaysia
e-mail: kharudin@uctati.edu.my

A. A. Z. Othman
e-mail: anwarzikri@kbc.com.my

J. K. S. Paw · C. T. Yaw
Institute of Sustainable Energy, Universiti Tenaga Nasional, The Energy University), Jalan Ikram-Uniten, Kajang, 43000 Selangor, Malaysia

J. D. Tan
School of Electrical Engineering and Artificial Intelligence, Xiamen University Malaysia, 43900 Sepang, Selangor, Malaysia

T. Yusaf
School of Engineering and Technology, Central Queensland University, Brisbane, QLD 4009, Australia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_18

brass by 17.12%. Thermal expansion, tied to bond energy and melting points, influences post-heating changes. Material thickness shifts due to expansion/contraction, while hardness alters due to microstructural shifts. Account for these changes in precise assessments. This effect was taken into account as a parameter relating with performance of thickness measurement.

Keywords Ultrasonic testing · Thickness · Hardness · Higher temperature

1 Introduction

Due to the significance of heat treatment verification, various evaluation methods have been developed. Among them are conventional destructive methods, including metallographic tests and assessments of mechanical properties through impact testing and tensile tests. Additionally, hardness measurement, mandated by many standards, is often employed to evaluate the quality of heat treatment [1–4]. The integrity of parts is irreversibly disrupted by destructive verification methods. Consequently, when producing pieces like gears, this method becomes unsuitable, limiting the assessment to surface hardness only. Clearly, this approach is inadequate for verifying treatment quality [5–13]. Although researchers have explored ultrasound for steel testing, no published studies examine the relationship between ultrasonic properties and grain hardness under different heat treatments. Some examinations of quenching processes in the fastener industry merely focus on hardness without metallographic analyses, and, in some cases, quenching hardness isn't even measured [14–17]. Ultrasonic thickness gauging extends beyond normal temperatures, allowing measurements even at surface temperatures reaching 500 °C. This technique enables measurements from one side of a hot part, making it favourable even when both sides are accessible. However, the effect of higher temperatures on sound velocity must be accounted for to ensure calibrated accuracy. While metals experience less sound velocity change than plastics, neglecting this correction may lead to significant errors in thickness measurement. Stainless Steel parts exhibit diverse characteristics, including microstructure, particle orientation, and heat treatment conditions, which can impact sound velocity during Ultrasonic Thickness Testing Method (UTTM). Hence, both traditional methods and ultrasonic processing can be used to examine and compare these conditions and environments [11, 18–25]. The gap lies in understanding post-heating material changes in thickness and hardness through ultrasonic testing (UT) in non-destructive testing (NDT). Justification arises from the need to assess structural integrity after heating. This gap necessitates research to establish correlations between ultrasonic responses and altered properties.

In response, this study aimed to alter the thickness and hardness material for Stainless steel, Brass steel and Alloy steel using varying quenching temperatures and observed the relationships between ultrasonic wave velocity, attenuation rate, grain size, and hardness at distinct frequencies (10 MHz) under different quenching temperatures. The results aim to offer fastener manufacturers a non-destructive alternative

for evaluating the heat treatment process to maintain product quality. Contribution includes enhancing post-heating evaluation accuracy. Sensory system significance lies in adapting UT for temperature-induced property variations and optimizing transducer coupling, ensuring precise measurements. Bridging this gap improves material assessment reliability, vital for industries relying on UT-based NDT for quality control and safety assurance in diverse thermal conditions. Non-Destructive Testing (NDT) proves to be a superior technique, as it allows assessment and monitoring of Stainless Steel structure thickness, condition, and specifications before or during procedures without causing damage to components or equipment. Furthermore, NDT eliminates the need to access both ends of longest pipe systems. This study also will assess the accuracy of UTTM for measuring thickness in Stainless Steel, Brass steel and Alloy steel materials overall.

2 Methodology

2.1 Ultrasonic Measurement Steps

This experiment required several process to be executed in order to get the thickness and hardness value. Initial setup required the specimen to heat up inside the post-weld machine into different temperature setup which is 100 °C until 500 °C. Upon reached certain temperature as required. Thickness measurement shall be measured and subsequently followed by hardness measurement. Both data shall be recorded, and data will be compared between ambient and elevated temperature. Following Fig. 1 Shown overall process of data measurement for this project.

The process of experiment start with calibration of ultrasonic thickness equipment (GE USM GO ultrasonic flow detection model). Calibration setup shall use equivalent material specification to attain similar velocities of material. For these purposes of experiment 3 type of material (Stainless steel, alloy, and brass) was prepared and calibration will take place according to the designated material. Calibration of ultrasonic equipment using standard step wedges. This calibration step wedges shall be certified and endorsed by third party body to ensure the accuracy of measurement. Figure 2a, b shown setup of calibration for material brass. Similar process applied for other type of materials. During this process same couplant shall be used for calibration and testing. Thickness measurement required couplant as a medium to transmit and transduce the ultrasonic energy/wave between the probe and test specimen. Usually the couplant in form of liquid. This experiment required high temperature couplant can withstand up to 500 °C. Couplant will typically be moderately viscous, nontoxic liquids, gels, or pastes. Figure 2c shown couplant applied on the specimen.

Prior heat up the specimen, the initial thickness will be measured at ambient temperature using zero-degree probe as show on Fig. 3. Four set of samples with three types of different material (Stainless steel, Brass steel and Alloy steel) were used for thickness measured.

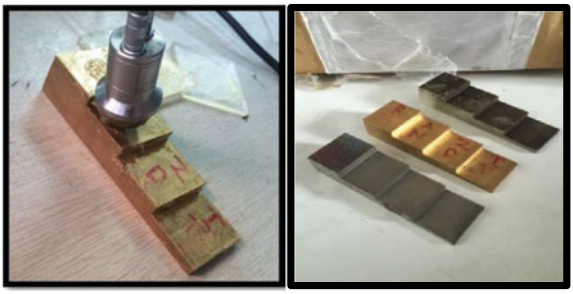


Fig. 1 Overall process flow chart

Upon completion of measured initial thickness, the next process is to heat up into specified temperature at 100, 200, 300, 400 and 500 °C. All plate will have remarked with symbol A, B, C and D. Figure 4 shown material being prepared for heat up. Final process upon reach of required temperature, four set of thickness measurement will be performed again to compare the thickness measured at ambient temperature. The reading also will total up and average accordingly. Figure 4 shown process of final thickness measurement.

2.2 Hardness Testing Method

Figure 5 shows the heat treatment process. The heat treatment machine Evenheat Kiln LB 22.5 Heat Treatment Oven (CE) 2023 Model was used for heat the metals until at certain temperature to alter their physical and mechanical properties without changing the product shape. Capable of all heat treating processes up to 2200 °F



(a) (b)



(c)

Fig. 2 **a** Calibration on brass step wedges, **b** three types of step wedges material and **c** couplant applying on the specimen

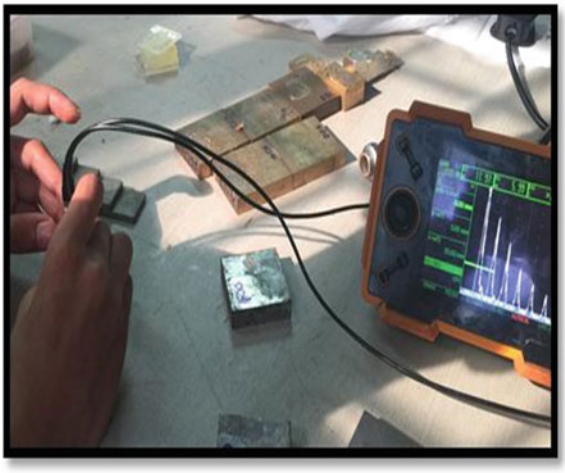


Fig. 3 Initial thickness measurement

(1204 °C). It's 571 mm deep design and power to safely ramp the product to 1000'C in approximately 1 h makes it a perfect fit for both private and low volume heat treaters. Heat treatment is sometimes done inadvertently due to manufacturing processes that either heat or cool the metal such as welding or forming.

Figure 6 shows hardness tester machine (RASNE-1). Major load 588.4, 980.7, 1471 N and minimum load is 98.07 N. Hardness is the resistance of a material to



(a) (b)

Fig. 4 **a** Specimen prepared for heat up process and **b** thickness measurement at elected temperature



Fig. 5 Heat treatment machine

localized deformation. The term can be applied to deformation from indentation, scratching, cutting or bending. In metals, ceramics and most polymers, the deformation considered is plastic deformation of the surface. The lack of a fundamental definition indicates that hardness is not a basic property of a material, but rather a composite one with contributions from the yield strength, work hardening, true tensile strength, modulus, and others factors. Hardness measurements are widely used for the quality control of materials because of its foot method and considered to be nondestructive tests when the marks or indentations produced by the test are in low stress areas.

Next process prior heat up to measure the hardness test. The hardness machine will be calibrated including the specimen will be measured the hardness at ambient



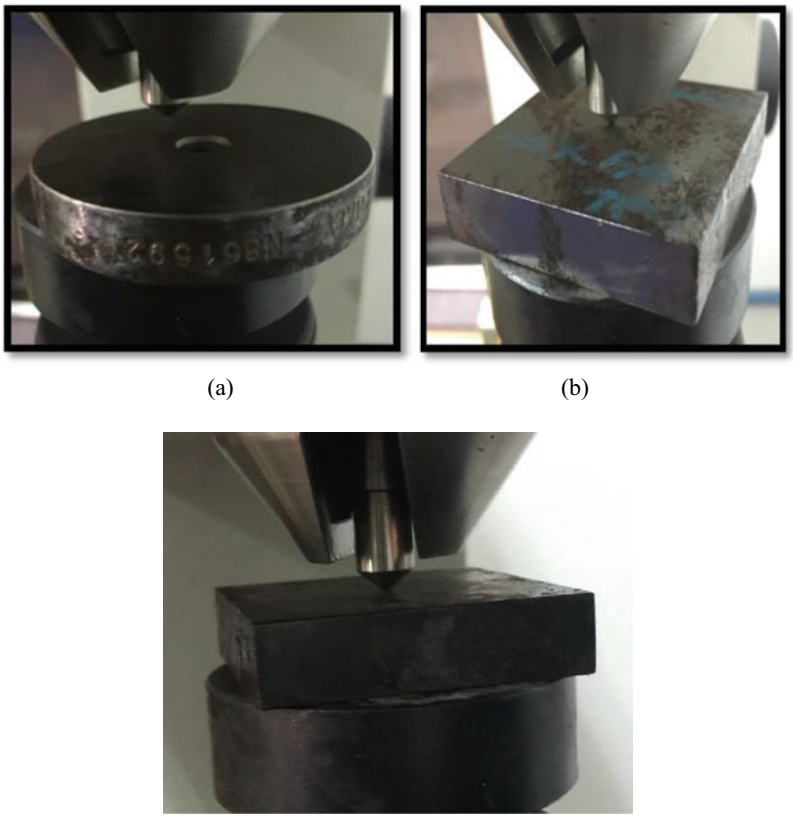
Fig. 6 Hardness tester machine

temperature at Fig. 7a, b. Baseline measurement is to ensure reading is accurate. Final process is to measure hardness after heating up into specified temperature. All data measured will be recorded accordingly for each specific temperature 100, 200, 300, 400 and 500 °C. Figure 7c showing process of hardness measurement for each sample.

3 Results

3.1 Thickness and Hardness Measurement for Stainless Steel

In this section the data from UT probes will be analyses by using Design of Expert software (DOE) for identify the quantity of sampling and also the significant data getting from experiment. From here the mathematical modelling will be generate with show the relation on each parameter used on this experiment. Table 1 presents the thickness measurements for stainless steel samples before and after the heating process. The measurement procedure involved recording the thickness values of the samples at various points. Sample A, which was initially at a temperature of 28 °C, had its thickness measured prior to the heating process. The measurements yielded values of 12.21 mm at points 1 and 2, with a subsequent drop to 12.20 mm at point 3, followed by an increase to 12.22 mm at point 4. The average thickness before heating was determined to be 12.21 mm. The next step involved subjecting sample A to heating at 100 °C, resulting in changes to its thickness. The thicknesses measured at points 1 and 2 were 12.26 mm, while at points 3 and 4, the thicknesses were 12.25 mm. Consequently, the average thickness of the stainless steel after being heated to 100 °C was calculated to be 12.26 mm. Sample B, which was also initially at 28 °C, underwent similar thickness measurements before and after heating. The measurements yielded values of 12.20 mm at point 1 and 12.22 mm at point 2 before heating. Subsequently, there was a drop to 12.21 mm at point 3, followed by an



(a) (b)

(c)

Fig. 7 **a** Calibration of hardness equipment, **b** sample calibrated at ambient temperature, **c** hardness measurement on the sample

increase to 12.23 mm at point 4. The average thickness before heating was determined to be 12.22 mm. After heating sample B to 200 °C, the thickness measurements were recorded as 12.45 mm at point 1, 12.40 mm at point 2, 12.54 mm at point 3, and 12.49 mm at point 4. The average thickness of the stainless steel after being heated to 200 °C was calculated to be 12.47 mm. This procedure was repeated for samples C, D, and E, ensuring that the thickness values before and after heating were accurately recorded, and subsequently, the average thicknesses were calculated for each sample.

Moving on to the hardness measurements for stainless steel samples, the results are also presented in Table 1. The initial step involved measuring the hardness of each sample. Sample A, initially at 28 °C, underwent hardness measurements, resulting in values of 34.20 at point 1 and 33.70 at point 2. At point 3, the hardness decreased to 33.50, and at point 4, it measured 31.80. The average hardness before heating was determined to be 33.30. Subsequently, sample A was heated to 100 °C, and the hardness values were recorded as 34.70 at point 1, 35.10 at point 2, 34.20 at point 3,

Table 1 Result of thickness and hardness measurement for stainless steel material

Sample	Temperature before heating (°C)	Thickness before heating (mm)	Hardness before heating	Temperature after heating (°C)	Thickness after heating (mm)	Hardness after heating (HRC)
A	28 (°C)	Point 1: 12.21	Point 1: 34.30	100 (°C)	Point 1: 12.26	Point 1: 34.70
		Point 2: 12.21	Point 2: 33.70		Point 2: 12.26	Point 2: 35.10
		Point 3: 12.20	Point 3: 33.50		Point 3: 12.25	Point 3: 34.20
		Point 4: 12.22	Point 4: 31.80		Point 4: 12.25	Point 4: 33.90
		Average: 12.21	**Average: 33.30**		**Average: 12.26**	**Average: 34.50**
B	28 (°C)	Point 1: 12.20	Point 1: 34.60	200 (°C)	Point 1: 12.45	Point 1: 35.30
		Point 2: 12.22	Point 2: 32.30		Point 2: 12.40	Point 2: 34.90
		Point 3: 12.21	Point 3: 32.60		Point 3: 12.54	Point 3: 34.80
		Point 4: 12.23	Point 4: 33.50		Point 4: 12.49	Point 4: 35.50
		Average: 12.22	**Average: 33.30**		**Average: 12.47**	**Average: 35.10**
C	28 (°C)	Point 1: 12.21	Point 1: 33.40	300 (°C)	Point 1: 12.58	Point 1: 37.60
		Point 2: 12.21	Point 2: 31.20		Point 2: 12.54	Point 2: 36.80
		Point 3: 12.20	Point 3: 31.50		Point 3: 12.55	Point 3: 37.30
		Point 4: 12.22	Point 4: 32.90		Point 4: 12.54	Point 4: 37.70
		Average: 12.21	**Average: 32.20**		**Average: 12.56**	**Average: 37.40**
D	28 (°C)	Point 1: 12.20	Point 1: 34.40	400 (°C)	Point 1: 12.61	Point 1: 38.20
		Point 2: 12.22	Point 2: 35.70		Point 2: 12.67	Point 2: 38.70
		Point 3: 12.21	Point 3: 35.20		Point 3: 12.65	Point 3: 37.70
		Point 4: 12.23	Point 4: 32.40		Point 4: 12.66	Point 4: 38.10
		Average: 12.22	**Average: 34.40**		**Average: 12.64**	**Average: 38.20**

(continued)

Table 1 (continued)

Sample	Temperature before heating (°C)	Thickness before heating (mm)	Hardness before heating	Temperature after heating (°C)	Thickness after heating (mm)	Hardness after heating (HRC)
E	28 (°C)	Point 1: 12.21	Point 1: 32.80	500 (°C)	Point 1: 12.88	Point 1: 40.20
		Point 2: 12.21	Point 2: 33.50		Point 2: 12.96	Point 2: 44.30
		Point 3: 12.20	Point 3: 33.40		Point 3: 12.79	Point 3: 39.80
		Point 4: 12.22	Point 4: 34.10		Point 4: 12.79	Point 4: 40.70
		Average: 12.21	**Average: 33.50**		**Average: 12.90**	**Average: 41.30**

and 33.90 at point 4. The average hardness of the stainless steel after being heated to 100 °C was calculated to be 34.50. For sample B, the hardness measurements before heating resulted in values of 34.60 at point 1 and 32.30 at point 2. At point 3, the hardness increased to 32.60, followed by a measurement of 33.50 at point 4. The average hardness before heating was determined to be 33.30. After heating sample B to 200 °C, the hardness values measured were 35.30 at point 1, 34.90 at point 2, 34.80 at point 3, and 35.50 at point 4. The average hardness of the stainless steel after being heated to 200 °C was calculated to be 35.10. Similar to the thickness measurements, the hardness measurements and corresponding heating temperatures were conducted for samples C, D, and E. The hardness values before and after heating were accurately recorded for each sample, and the average harnesses were calculated accordingly.

The significance of the model is implied by the Model F-value of 53.31. It is highly improbable for a "Model F-Value" of this magnitude to arise solely from noise, with only a 0.01% chance. Model terms are considered significant when the "Prob > F" values are less than 0.0500. In this case, B $++2+-$ represents a significant model term. Conversely, model terms are deemed insignificant when the "Prob > F" values exceed 0.1000. If there are numerous insignificant model terms (excluding those needed to support hierarchy), reducing the model might lead to an improvement. Table 2 and Fig. 8 show the relation of heating, thickness and temperature for Stainless Steel material.

Table 2 ANOVA response surface quadratic for stainless steel

ANOVA for response surface quadratic. Model analysis of variance table [partial sum of squares—Type III]						
Source	Sum of Squares	df	Mean Square	F Value	*p*-value Prob > F	
Model	88.95177477	5	17.79035495	53.31202102	<0.0001	Significant
A-heating	0.244940286	1	0.244940286	0.734007933	0.4199	
B-thickness	1.711215174	1	1.711215174	5.127966223	0.0579	
AB	1.196583811	1	1.196583811	3.585780129	0.1001	
A^2	0.672458917	1	0.672458917	2.015144949	0.1987	
B^2	2.015246078	1	2.015246078	6.039049898	0.0436	
Residual	2.335917534	7	0.333702505			
Cor total	91.28769231	12				

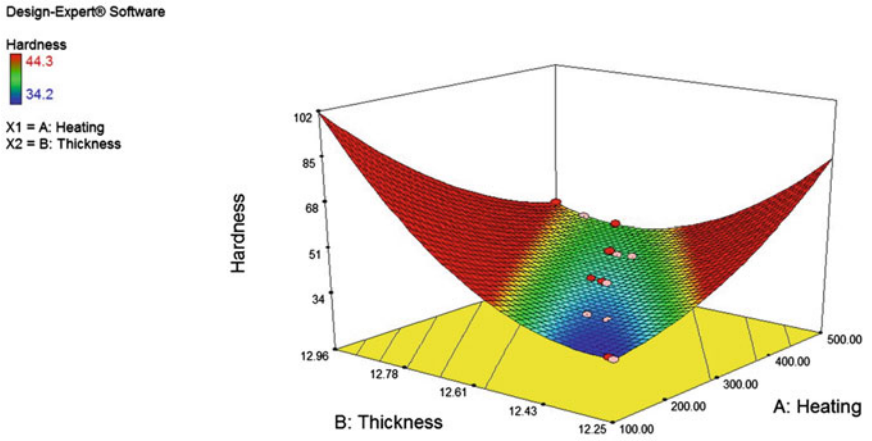


Fig. 8 Relation for heating, thickness and hardness for stainless steel

Final Equation in Terms of Actual Factors:

$$\begin{aligned} Hardness = {} & (21584.306935684) + (4.4513522140596 * Heating) \\ & - (3546.0398180757 * Thickness) \\ & - (0.36511234960453 * Heating * Thickness) \\ & + \left(0.00022821953793355 * Heating^2\right) \\ & + \left(145.86469051742 * Thickness^2\right) \end{aligned} \quad (1)$$

3.2 Thickness and Hardness Measurement for Brass

In this section, the focus on the thickness and hardness measurements for brass samples. The results are presented in Table 3. The procedure involved measuring the thickness of each sample before and after the heating process. Sample A, initially at a temperature of 28 °C, underwent thickness measurements before the heating process. The recorded values were 12.21 mm at points 1 and 2, with a drop to 12.20 mm at point 3 and an increase to 12.22 mm at point 4. The average thickness before heating was determined to be 12.21 mm. Subsequently, sample A was heated to 100 °C, resulting in changes to its thickness. The measured thicknesses were 12.74 mm at point 1, 12.69 mm at point 2, 12.70 mm at point 3, and 12.69 mm at point 4. The average thickness of the brass after being heated to 100 °C was calculated to be 12.71 mm.

Sample B, also initially at 28 °C, underwent similar thickness measurements. The measurements yielded values of 12.20 mm at point 1 and 12.22 mm at point 2 before heating. At point 3, the thickness dropped to 12.21 mm, followed by an increase to 12.23 mm at point 4. The average thickness before heating was determined to be 12.22 mm. After heating sample B to 200 °C, the thickness measurements were recorded as 12.84 mm at point 1, 12.85 mm at point 2, 12.81 mm at point 3, and 12.88 mm at point 4. The average thickness of the brass after being heated to 200 °C was calculated to be 12.85 mm. Similar to the previous section, the thickness measurements and corresponding heating temperatures were carried out for samples C, D, and E. The thickness values before and after heating were recorded accurately for each sample, and the average thicknesses were calculated accordingly.

Transitioning to the hardness measurements, Table 3 also displays the results for the hardness of brass samples before and after the heating process. The initial step involved measuring the hardness of each sample. For sample A, at an initial temperature of 28 °C, the hardness was measured and resulted in values of 22.30 at point 1 and 24.60 at point 2. At point 3, the hardness decreased to 24.60, and at point 4, it was 22.30. The average hardness before heating was determined to be 23.20. After heating the sample to 100 °C, the hardness values were recorded as 23.40 at point 1, 23.70 at point 2, 24.20 at point 3, and 23.90 at point 4. The average hardness of the brass after being heated to 100 °C was calculated to be 23.80. Similarly, for sample B, initially at 28 °C, the hardness measurements before heating resulted in values of 24.30 at point 1 and 21.80 at point 2. At point 3, the hardness increased to 23.50, followed by a measurement of 23.70 at point 4. The average hardness before heating was determined to be 23.30. After heating sample B to 200 °C.

The model's significance is reflected by the Model F-value of 23.61. The likelihood of a "Model F-Value" of this magnitude occurring due to noise is merely 0.02%. Model terms achieve significance when the "Prob > F" values are less than 0.0500. In this case, significant model terms include B and B $++2$ $+-$. Conversely, model terms are considered not significant when the "Prob > F" values exceed 0.1000. If there are numerous model terms lacking significance (excluding those required to support hierarchy), your model may be improved through model reduction. Table 4

Table 3 Result of thickness and hardness measurement brass steel material

Sample	Temperature before heating (°C)	Thickness before heating (mm)	Hardness before heating	Temperature after heating (°C)	Thickness after heating (mm)	Hardness after heating (HRC)
A	28 (°C)	Point 1: 12.21	Point 1: 22.30	100 (°C)	Point 1: 12.74	Point 1: 23.40
		Point 2: 12.21	Point 2: 24.60		Point 2: 12.69	Point 2: 23.70
		Point 3: 12.20	Point 3: 23.50		Point 3: 12.70	Point 3: 24.20
		Point 4: 12.22	Point 4: 22.30		Point 4: 12.69	Point 4: 23.90
		Average: 12.21	**Average: 23.20**		**Average: 12.71**	**Average: 23.80**
B	28 (°C)	Point 1: 12.20	Point 1: 24.30	200 (°C)	Point 1: 12.84	Point 1: 24.10
		Point 2: 12.22	Point 2: 21.80		Point 2: 12.85	Point 2: 24.70
		Point 3: 12.21	Point 3: 23.50		Point 3: 12.81	Point 3: 23.90
		Point 4: 12.23	Point 4: 23.70		Point 4: 12.88	Point 4: 24.50
		Average: 12.22	**Average: 23.30**		**Average: 12.85**	**Average: 24.30**
C	28 (°C)	Point 1: 12.21	Point 1: 24.40	300 (°C)	Point 1: 13.33	Point 1: 24.80
		Point 2: 12.21	Point 2: 25.70		Point 2: 13.28	Point 2: 25.10
		Point 3: 12.20	Point 3: 22.40		Point 3: 13.31	Point 3: 25.70
		Point 4: 12.22	Point 4: 25.20		Point 4: 13.38	Point 4: 25.30
		Average: 12.21	**Average: 24.40**		**Average: 13.33**	**Average: 25.20**
D	28 (°C)	Point 1: 12.20	Point 1: 24.10	400 (°C)	Point 1: 13.65	Point 1: 25.30
		Point 2: 12.22	Point 2: 23.50		Point 2: 13.70	Point 2: 25.70
		Point 3: 12.21	Point 3: 22.80		Point 3: 13.73	Point 3: 24.80
		Point 4: 12.23	Point 4: 23.40		Point 4: 13.71	Point 4: 25.10
		Average: 12.22	**Average: 23.50**		**Average: 13.70**	**Average: 25.20**

(continued)

Table 3 (continued)

Sample	Temperature before heating (°C)	Thickness before heating (mm)	Hardness before heating	Temperature after heating (°C)	Thickness after heating (mm)	Hardness after heating (HRC)
E	28 (°C)	Point 1: 12.21	Point 1: 22.90	500 (°C)	Point 1: 13.82	Point 1: 25.80
		Point 2: 12.21	Point 2: 23.40		Point 2: 13.83	Point 2: 26.10
		Point 3: 12.20	Point 3: 21.50		Point 3: 13.81	Point 3: 26.30
		Point 4: 12.22	Point 4: 21.20		Point 4: 13.82	Point 4: 25.90
		Average: 12.21	**Average: 22.20**		**Average: 13.82**	**Average: 26.00**

and Fig. 9 show the relation between Hardness, Thickness and Temperature for heating process for Brass steel.

Final Equation in Terms of Coded Factors:

$$Hardness = (+22.03) + (3.35 * B) + (0.33 * A * B) + (1.21 * A2) - (0.83 * B2) \quad (2)$$

Table 4 ANOVA response surface quadratic for brass steel

ANOVA for response surface quadratic. Model analysis of variance table [partial sum of squares—Type III]						
Source	Sum of squares	df	Mean Square	F Value	*p*-value Prob > F	
Model	8.318376122	4	2.07959403	23.60824936	0.0002	Significant
B-thickness	1.608204853	1	1.608204853	18.25688122	0.0027	
AB	0.09340611	1	0.09340611	1.060377504	0.3333	
A^2	0.362865615	1	0.362865615	4.119372245	0.0769	
B^2	1.17404773	1	1.17404773	13.32818385	0.0065	
Residual	0.704700801	8	0.0880876			
Cor total	9.023076923	12				

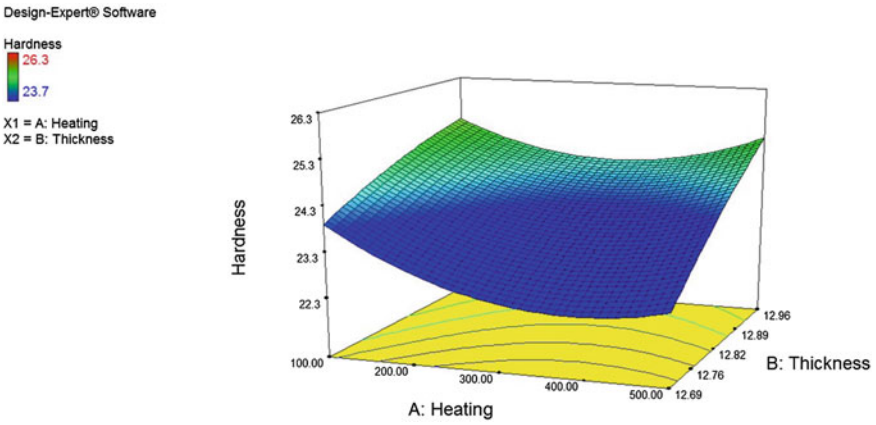


Fig. 9 Relation for heating, thickness and hardness for brass steel

3.3 Thickness and Hardness Measurement for Alloy Steel

Alloy steel, the final material chosen for UT inspection in this experiment, had its properties thoroughly examined. Sample A, with an initial temperature of 28 °C, underwent precise measurements to determine its thickness before the heating process. The alloy steel displayed a thickness of 12.20 mm at point 1 and 12.22 mm at point 2. At point 3, a decrease in thickness was observed, reaching 12.21 mm, followed by an increase to 12.23 mm at point 4. By employing meticulous techniques, the average thickness before heating was determined to be 12.22 mm. Subsequently, sample A was subjected to a temperature of 100 °C, inducing notable alterations in its thickness. The recorded measurements yielded values of 12.55 mm at point 1, 12.60 mm at point 2, 12.59 mm at point 3, and 12.57 mm at point 4. Through careful calculations, the average thickness of the alloy steel after being heated to 100 °C was determined to be 12.57 mm. The comprehensive results, highlighting the thickness measurements for alloy steel samples before and after the heating process, can be found in Table 5.

Similarly, sample B, which shared the same initial temperature of 28 °C, underwent meticulous thickness measurements before heating. The recorded measurements exhibited values of 12.21 mm at both points 1 and 2. At point 3, the thickness experienced a decline, reaching 12.20 mm, followed by an increase to 12.22 mm at point 4. The precise calculations established the average thickness before heating at 12.21 mm. Upon exposure to a temperature of 200 °C, sample B underwent changes in thickness, as indicated by measurements of 12.86 mm at point 1, 12.89 mm at point 2, 12.85 mm at point 3, and 12.86 mm at point 4. Through meticulous analysis, the average thickness of the alloy steel after being heated to 200 °C was determined to be 12.86 mm.

It similar for samples C, D, and E. Each sample underwent precise measurements of thickness before and after the heating process, ensuring accurate recordings of the

Table 5 Result of thickness and hardness measurement alloy steel material

Sample	Temperature before heating (°C)	Thickness before heating (mm)	Hardness before heating	Temperature after heating (°C)	Thickness after heating (mm)	Hardness after heating (HRC)
A	28 (°C)	Point 1: 12.20	Point 1: 32.80	100 (°C)	Point 1:12.55	Point 1: 33.60
		Point 2: 12.22	Point 2: 33.50		Point 2:12.60	Point 2: 32.90
		Point 3: 12.21	Point 3: 33.40		Point 3:12.59	Point 3: 35.70
		Point 4: 12.23	Point 4: 34.10		Point 4:12.57	Point 4: 35.40
		Average: 12.22	**Average: 33.50**		**Average: 12.57**	**Average: 34.40**
B	28 (°C)	Point 1: 12.21	Point 1: 34.40	200 (°C)	Point 1:12.86	Point 1: 35.70
		Point 2: 12.21	Point 2: 33.70		Point 2:12.89	Point 2: 35.40
		Point 3: 12.20	Point 3: 31.80		Point 3:12.85	Point 3: 34.90
		Point 4: 12.22	Point 4: 33.50		Point 4:12.86	Point 4: 35.50
		Average: 12.21	**Average: 33.30**		**Average: 12.86**	**Average: 35.57**
C	28 (°C)	Point 1: 12.20	Point 1: 34.60	300 (°C)	Point 1: 13.28	Point 1: 37.70
		Point 2: 12.22	Point 2: 32.30		Point 2: 13.31	Point 2: 37.90
		Point 3: 12.21	Point 3: 32.60		Point 3:13.33	Point 3: 36.50
		Point 4: 12.23	Point 4: 32.90		Point 4: 13.28	Point 4: 37.20
		Average: 12.22	**Average: 33.10**		**Average: 13.30**	**Average: 37.32**
D	28 (°C)	Point 1: 12.21	Point 1: 34.40	400 (°C)	Point 1:13.49	Point 1: 38.70
		Point 2: 12.21	Point 2: 35.10		Point 2:13.46	Point 2: 38.50
		Point 3: 12.20	Point 3: 35.30		Point 3:13.43	Point 3: 37.80
		Point 4: 12.22	Point 4: 32.40		Point 4:13.42	Point 4: 38.90
		Average: 12.21	**Average: 34.40**		**Average: 13.45**	**Average: 38.47**

(continued)

Table 5 (continued)

Sample	Temperature before heating (°C)	Thickness before heating (mm)	Hardness before heating	Temperature after heating (°C)	Thickness after heating (mm)	Hardness after heating (HRC)
E	28 (°C)	Point 1: 12.20	Point 1: 32.90	500 (°C)	Point 1:14.05	Point 1: 39.90
		Point 2: 12.22	Point 2: 33.40		Point 2:14.03	Point 2: 40.10
		Point 3: 12.21	Point 3: 33.60		Point 3:14.01	Point 3: 44.30
		Point 4: 12.23	Point 4: 34.20		Point 4:14.03	Point 4: 41.30
		Average: 12.22	**Average: 33.50**		**Average: 14.03**	**Average: 41.40**

values. The average thicknesses were then meticulously calculated, allowing for a comprehensive comparison between the different samples.

Furthermore, Table 5 presents a comprehensive overview of the hardness measurements conducted on the alloy steel samples before and after the heating process. Sample A, with an initial temperature of 28 °C, had its hardness accurately measured, resulting in values of 32.80 at point 1 and 33.50 at point 2. At point 3, a decline in hardness was observed, reaching 33.40, followed by an increase to 34.10 at point 4. The meticulous determination of the average hardness before heating yielded a value of 33.50. Subsequently, the sample was exposed to a temperature of 100 °C, leading to the recording of hardness values at 33.60 for point 1, 32.90 for point 2, 35.70 for point 3, and 35.40 for point 4. Through precise calculations, the average hardness of the alloy steel after being heated to 100 °C was determined to be 34.40.

Similar to the thickness measurements, the hardness measurements and corresponding heating temperatures were conducted with utmost precision for samples C, D, and E. Accurate recordings of hardness values before and after heating were ensured for each sample, allowing for meticulous calculations of the average harnesses.

The model's significance is evidenced by the Model F-value of 46.41. The occurrence of a "Model F-Value" of this magnitude due to noise is associated with a mere 0.01% probability. Model terms are considered significant when the "Prob > F" values are less than 0.0500. In this case, significant model terms are represented by B. Conversely, model terms are deemed not significant when the "Prob > F" values exceed 0.1000. Should there be numerous model terms lacking significance (excluding those required to support hierarchy), model reduction might be employed to enhance your model. Table 6 and Fig. 10 show the relation between Hardness, Thickness and Temperature for heating process for Alloy steel.

Final Equation in Terms of Actual Factors:

Table 6 ANOVA response surface quadratic for alloy steel

ANOVA for response surface reduced linear model. Analysis of variance table [partial sum of squares—Type III]						
Source	Sum of squares	df	Mean Square	F Value	*p*-value Prob > F	
Model	74.6575317	1	74.65753166	46.4110338	<0.0001	Significant
B-thickness	74.6575317	1	74.65753166	46.4110338	<0.0001	
Residual	17.694776	11	1.608616003			
Cor total	92.3523077	12				

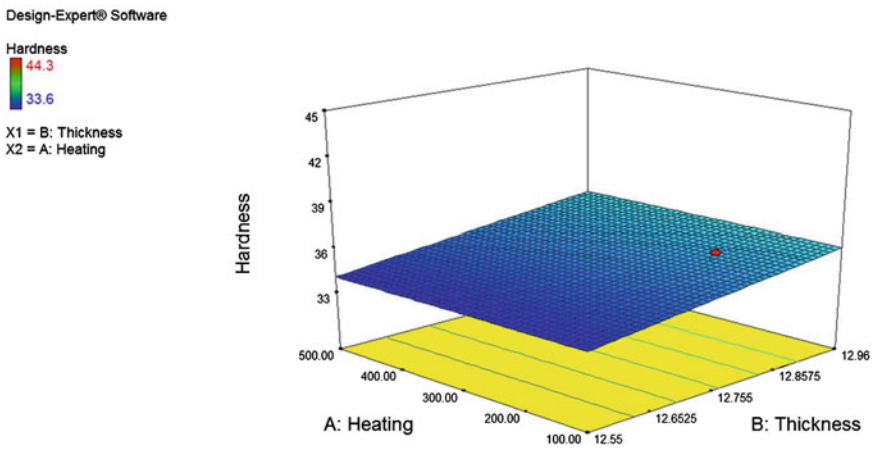


Fig. 10 Relation for Heating, Thickness and Hardness for Alloy Steel

$$Hardness = (-26.89720) + (4.85979 * Thickness) \quad (3)$$

4 Discussion

The data in Table 1 presents measurements of thickness and hardness for stainless steel samples before and after heating at different temperatures. Samples A, B, C, D, and E were initially at 28 °C, and their properties were recorded at multiple points. Heating resulted in varying effects on thickness and hardness. Sample A showed a slight increase in both thickness and hardness after heating to 100 °C, while Sample B exhibited more significant changes after heating to 200 °C. The findings suggest that the impact of heating on stainless steel properties is temperature-dependent. While some samples displayed subtle changes, others experienced more pronounced alterations. These insights can be valuable for material selection and design in specific

applications, aiding materials scientists and engineers in understanding stainless steel behavior under different thermal conditions. The discussion in Table 3 focuses on the thickness and hardness measurements of brass samples before and after the heating process. For Sample A, heating to 100 °C resulted in an average thickness increase from 12.21 mm to 12.71 mm and an average hardness increase from 23.20 to 23.80. Similarly, Sample B exhibited changes after heating to 200 °C, with the average thickness increasing from 12.22 to 12.85 mm and the average hardness rising from 23.30 to 23.70. The data suggests that the heating temperature had an impact on both thickness and hardness for the brass samples. As seen in the results for samples A and B, higher heating temperatures led to more significant changes in both properties. This information can be crucial for understanding how brass responds to thermal conditions and can aid in material selection and design for applications where temperature fluctuations are critical. Further investigations for samples C, D, and E can provide a comprehensive understanding of how brass behaves under different heating scenarios.

The chosen material for ultrasonic testing (UT) inspection in this experiment was alloy steel. The thickness measurements of Sample A were meticulously performed, revealing an average thickness of 12.22 mm before heating. Upon heating the sample to 100 °C, notable changes were observed, and the average thickness increased to 12.57 mm. Similarly, Sample B exhibited a rise in average thickness from 12.21 to 12.86 mm after being subjected to 200 °C. The hardness measurements for the alloy steel samples were also conducted with utmost precision. Sample A had an average hardness of 33.50 before heating, which increased to 34.40 after being exposed to 100 °C. The same level of accuracy was applied to Samples C, D, and E, allowing for thorough comparisons of the thickness and hardness variations at different heating temperatures for the alloy steel samples. Table 5 presents a comprehensive overview of the results, providing valuable insights for the behavior of alloy steel under varying thermal conditions. The precise measurements of thickness and hardness contribute to a better understanding of the material's response to temperature changes, and this data can be essential in guiding material selection and design for practical applications.

In the study, ultrasonic testing (UT) was employed for assessing material thickness and hardness after a heating process. For stainless steel, samples showed varying thickness and hardness changes post-heating. Similar trends were noted in brass and alloy steel samples, confirming that heating induces alterations.

As the comparison for thickness measurement the aluminum, Maximum thinning is 18.28% at (200 C) less than initial thickness in warm deep drawing and 25.08% less than initial thickness in deep drawing at room temperature it means that the increasing of thickness is 6.8% [26]. On experiment the increasing for stainless steel is 2% in each 100 C degree increasing. It also supports actual UT inspection where, in steel this velocity change is approximately 1% per 55 °C or 100 °F change in temperature. (The exact value varies depending on the alloy.) In plastics and other polymers, this change is much greater, and can approach 50% per 55 °C or 100 °F change in temperature up to the melting point. If a temperature/velocity plot for the material is not available, then a velocity calibration should be performed on a

sample of the test material at the actual test temperature [27]. On here the data were getting on this experiment more accurate comparing the previous researcher based on supporting the new paper on UT method [26]. However, limitations such as a small sample size and a restricted temperature range might restrict generalizability.

The advantages of this research proven quantitative data were used for evaluation with accuracy, cost effective, ease of calibration, practical and without inducing damage to the material being tested. Socially, this research has practical implications. Industries reliant on materials subjected to heating processes, such as manufacturing and construction, could benefit from enhanced UT techniques. Accurate post-heating assessments could lead to better quality control, safer products, and optimized processes, fostering improved industrial performance and public safety. Moreover, the development of specialized sensory systems within UT could open doors to more effective and comprehensive non-destructive evaluations, positively impacting various sectors and ensuring the longevity and safety of critical infrastructure. This research has its limitation particularly on how to control the temperature during the measurement was taken.

5 Conclusion and Recommendation

In conclusion this research has demonstrate the effect of temperature in measurement of thickness and hardness, method of non-destructive testing (NDT) serve as crucial tools in various scientific and industrial domains, enabling the evaluation of material properties without inducing damage. NDT finds applications across diverse industries, with constant advancements leading to the development of novel techniques and applications. While NDT is commonly employed to assess the remaining wall thickness of objects vulnerable to corrosion or erosion, its utility extends to scenarios involving elevated temperatures that challenge conventional measurement approaches. Specifically, this study has focused on analyzing the changes in material thickness and hardness before and after heating processes, involving three distinct materials: stainless steel, alloy, and brass. The heating process was conducted within a temperature range of 100–500 °C. The findings underscore the significant impact of heating on material properties. Notably, the alloy steel exhibited the highest increase in thickness, with a substantial rise of up to 14.1%, followed by brass at 13.8% and stainless steel at 5.65%. In terms of hardness, stainless steel demonstrated the most notable increase, registering an uptick of 23.8%, followed closely by alloy steel at 23.58%, and brass at 17.12%. These results illuminate the complex interplay between temperature and material characteristics, underscoring the significance of thermal effects on material behavior. It's worth noting that thermal expansion, influenced by bond energy and melting points, exhibits a general trend of decrease with increasing bond energy. Given these insights, it is recommended that industries and researchers alike consider the implications of heating on material properties when assessing structural integrity and performance. Additionally, the findings emphasize

the need for tailored non-destructive testing methods that account for temperature-induced variations, particularly in scenarios where conventional measurements might not suffice. This research should be further developed with input from industries to enhance Ultrasonic Thickness Measurement focusing on elevated temperature.

Acknowledgements Our sincere appreciation goes to University College TATI for their invaluable support in providing the necessary facilities, expertise, and equipment for this experimental work. The assistance received from the Electrical Lab, Sensor Technology Lab, Electronic Lab, and Microcontroller Lab was crucial in conducting the research successfully.

References

1. Rifai D, Ibrahim MF, Ali K, Abdalla AN, Faraj MA, Pazikadin AR (2021) Investigation the effect of heat treatment on brass defect measurement using Eddy current testing. J Phys Conf Ser 1874. https://doi.org/10.1088/1742-6596/1874/1/012060
2. Segreto T, Bottillo A, Palmieri B, Nele L, Teti R (2020) Ultrasonic evaluation of induction heat treatment applied to thermoplastic matrix CFRP. Procedia CIRP 88:467–472. https://doi.org/10.1016/j.procir.2020.05.081
3. Ullah R, Fangnon E, Huuki J (2023) Effect of ultrasonic burnishing parameters on burnished-surface quality of stainless steel after heat treatment. Springer International Publishing. https://doi.org/10.1007/978-3-031-18326-3_4
4. Ali K, Joraimee Mohamad A, Rifai D, Paw KS, Phing CC, Yaw CT, Ding TJ (2022) Control, instrumentation and mechatronics: theory and practice. LNEE, vol 921, p 866. https://doi.org/10.1007/978-981-19-3923-5
5. Wu R, Zhang H, Yang R, Chen W, Chen G (2021) Nondestructive testing for corrosion evaluation of metal under coating. J Sens 2021. https://doi.org/10.1155/2021/6640406
6. Fuštar B, Lukačević I, Dujmović D (2018) Review of fatigue assessment methods for welded steel structures. Adv Civ Eng 2018. https://doi.org/10.1155/2018/3597356
7. Beneš P, Vítovec V (2021) Selected NDT techniques in heat treatment control. Manuf Technol 21:20–28. https://doi.org/10.21062/mft.2021.007
8. Ji B, Zhang H, Cao J, Zhang Q (2023) Evaluation of the thickness of each layer of Cu/Al laminate using laser ultrasonic. Coatings 13. https://doi.org/10.3390/coatings13030645
9. Mohamad AJ, Ali K, Rifai D, Salleh Z, Othman AAZ (2023) Eddy current testing methods and design for pipeline inspection system: a review. J Phys Conf Ser 2467. https://doi.org/10.1088/1742-6596/2467/1/012030
10. Ali K, Joraimee Mohamad A, Rifai D, Hazim Aizuddin Othman M, Mohamed Ayub MA, Hussian AH, Paw KS, Phing CC, Yaw CT, Ding TJ (2022) Investigation of GMR Eddy current permanent magnet system for carbon steel corrosion identification using IoT unmanned aerial vehicles drone platform. J Phys Conf Ser 2319. https://doi.org/10.1088/1742-6596/2319/1/012001
11. Ali K, Paw JKS, Rifai D, Awang NA, AdbAlla AN, Pazikadin AR, Hen CKH (2021) Long range ultrasonic testing system based on lamb wave method for validation an optimized piezoelectric sensor gap array. Springer Nat 1, 205–216 (2021). https://doi.org/10.1007/978-981-33-6926-9_19
12. Thornicroft K, Mares C, Mudge P (2012) Time-frequency analysis of long range ultrasonic signals. J Phys Conf Ser 382. https://doi.org/10.1088/1742-6596/382/1/012060
13. Ali K, Paw JKS, Rifai D, Abda Alla AN, Faraj MA, Hen CK (2020) Optimization of ultrasonic testing for defect localization signal on S55C carbon steel pipe. Int J Eng Sci Comput IJESC 10:24837–24842

14. Faraj MA, Samsuri FB, Abdalla AN, Rifai D, Ali K, Samsuri FB, Rifai D, Ali K (2018) Construct coil probe using GMR sensor for Eddy current testing. MATEC Web Conf 225:1–8. https://doi.org/10.1051/matecconf/201822502021
15. Ali K, Paw JKs, Aizat M, Sulaiman M, Abdalla AN, Hen CK (2018) S55c carbon steel for heat treatment process with different medium in attenuation measurement using ultrasonic testing. Int J Eng Technol 5:527–533
16. Hsia SY, Chou YT (2015) Assessing the hardness of quenched medium steel using an ultrasonic nondestructive method. Adv Mater Sci Eng 2015. https://doi.org/10.1155/2015/684836
17. Ali K, Rifai D, Paw JKs, Abdalla AN, Hen CK, Sukor MFM, Ismail Z, Ding TJ (2020) Effect of crack on bending process for S55c carbon steel with ultrasonic testing on zero degree probes. J Phys Conf Ser 1532. https://doi.org/10.1088/1742-6596/1532/1/012027
18. Zulkipli NA, Yusup EM (2021) The evaluation for accuracy of non-destructive testing (NDT) in ultrasonic inspection on mild steel material by ultrasonic testing thickness measurement (UTTM). Res Prog Mech Manuf Eng 2:41–47
19. Lv F, Zhang P, Tang Z, Yue Y, Yang K (2019) A guided wave transducer with sprayed magnetostrictive powder coating for monitoring of aluminum conductor steel-reinforced cables. Sensors (Switzerland). 19:1–15. https://doi.org/10.3390/s19071550
20. Bernieri A, Ferrigno L, Laracca M, Rasile A, Ricci M (2017) Ultrasonic non destructive testing on aluminium forged bars. In: IEEE international workshop on metrology for aerospace (MetroAeroSpace). IEEE, Padua, Italy, pp 223–227
21. Ali K, Mohamad AJ, Othman AAZ, Rifai D, Salleh Z, Paw JKS, Yaw CT, Tan JD (2023) Effect of metallic coating thickness on stainless steel and carbon steel welded plate using Eddy current testing. J Phys Conf Ser 2467. https://doi.org/10.1088/1742-6596/2467/1/012031
22. Ali K (2020) Piezoelectric sensor array for detecting gas pipeline crack. http://library.uniten.edu.my/
23. Abdalla AN, Ali K, Koh J, Paw S, Hen CK, Ding TJ (2018) Crack identification using piezoelectric testing on carbon steel pipe for transverse, longitudinal and hole defects with low excitation frequency. Int J Eng Technol 7:171–176
24. Abdalla AN, Ali K, Paw JKS, Rifai D, Faraj MA (2018) A novel eddy current testing error compensation technique based on Mamdani-type fuzzy coupled differential and absolute probes. Sensors (Switzerland) 18. https://doi.org/10.3390/s18072108
25. Rifai D, Pazikadin AR, Ali K, Faraj MA, Khamsah N, Abdalla AN (2020) Fuzzy logic error compensation scheme for Eddy current testing measurement on mild steel super facial crack. LNEE, vol 746. Springer, pp 225–243
26. Basril MAM, Teng HM, Azuddin M, Choudhury IA (2017) The effect of heating temperature and methods towards the formability of deep drawn square metal cup. IOP Conf Ser Mater Sci Eng 210. https://doi.org/10.1088/1757-899X/210/1/012067
27. Li M, He J (2021) Effect of high temperature on ultrasonic velocity in graphite. Diam Relat Mater 116:1–5. https://doi.org/10.1016/j.diamond.2021.108368

Enhancing Infant Safety: Performance Analysis of Deep Learning Method on Development Board for Real-Time Monitoring



Nugroho Budi Prasetyo, Dien Rahmawati, Wahmisari Priharti, and Muhammad Dhalhaz

Abstract Supervising an infant while managing household tasks poses a significant challenge for parents, often leading to safety concerns due to limited time and attention for constant monitoring. Conventional sensor-based monitoring systems have limitations in accurately detecting an infant's condition and position. This research proposes a deep learning algorithm method as an alternative to sensors. The deep learning algorithm will be integrated with a camera to assess the infant's position and condition using the Convolutional Neural Network (CNN) method. To enhance performance, a development board integrated with a Video Graphic Array (VGA) is utilized to accelerate the processing time of the deep learning method, and the CNN model is simplified using the MobileNetv2 architecture to reduce model weight. Testing with various scenarios, including an infant in potentially hazardous positions (i.e., prone, covered by an obstacle, or standing near a fence) and an infant in safe positions, achieved an accuracy of 94%. This indicates that the model can effectively identify the infant's position and determine their well-being. Additionally, the Frame Per Second (FPS) results of 20 demonstrate real-time infant supervision capabilities, allowing for timely intervention in case of any potential risks. Based on the experimental findings, this method holds promise for applications in the medical field and for parents who may struggle with continuous infant supervision, enabling them to fulfill other responsibilities without compromising infant safety.

Keywords Deep learning · Convolutional neural network · MobileNetv2 · TensorRT

N. B. Prasetyo · D. Rahmawati (✉) · W. Priharti · M. Dhalhaz
School of Electrical Engineering, Telkom University, Bandung, Indonesia
e-mail: dienrahmawati@telkomuniversity.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_19

1 Introduction

Ensuring the safety of infants, especially in their early stages of development is a primary concern for parents [1]. Households that are constantly in demand of other responsibilities often left infants unsupervised and end up risking their safety. Sudden movements such as thumb-sucking, gesturing, sneezing, etc., are integral to their gross motor development [2, 3]. However, without proper supervision, these can potentially place infants in precarious situations, leading to suffocation or harm from nearby objects, and in severe cases, resulting in Sudden Infant Death Syndrome (SIDS) [4].

While exhaustive inquiries, including thorough medical history assessments, have been conducted in an attempt to unravel the underlying causes of Sudden Infant Death Syndrome (SIDS), the precise etiology remains elusive [5]. Recent advancements in technology have ushered in a multitude of potential solutions for enhancing infant safety through the implementation of AI-integrated camera systems. These systems, in their current state, provide the ability to identify potential risks or irregularities during infancy without the need for constant supervision. Nevertheless, it is essential to note that the current technological landscape primarily focuses on minimizing the impact of incidents once they have already occurred, rather than proactively preventing them.

Existing studies have largely focused on the basic aspects of infant monitoring, such as recognizing the presence of an infant or analyzing their cries, rather than addressing more in-depth SIDS [6–11]. The research that has been carried out may be good enough to apply to some cases that are not too complex, and the resulting tool is also more practical. However, these studies often rely on conventional computing systems, thereby limiting their portability and scalability [12], such as only detecting whether the baby is present or not and only detecting their presence, not paying attention to the baby's sleeping position. The results of research conducted in [13] show that computer vision is unable to classify whether the baby's position is safe or not. Which is the dominant factor in the occurrence of SIDS, especially in babies aged 0 to 1 year.

Given these limitations, there is a need to develop a more comprehensive and effective approach to infant safety that utilizes state-of-the-art technology and addresses the important issue of SIDS prevention. This research seeks to bridge this gap by exploiting the potential of AI-integrated camera systems, specifically by integrating Convolutional Neural Network (CNN) into the system. This study aimed to improve on those systems to more accurately detect SIDS by employing position-based AI-integrated camera systems, thus offering a holistic approach to preventing SIDS. Such a system could only be achieved by using an efficient and resource-friendly approach, one of them is to use CNN-based models to handle image recognition problems by extracting important features of image models in these systems. impedes their practicality on resource-constrained development boards, limiting their scalability and efficacy. CNN uses these convolution layers to highlight a specific feature of an image with a specific filter so it can be identified more easily regardless of the

environment to detect potential SIDS. Hopefully, the addition of CNN in real-time settings could address the pivotal issue of infant safety, particularly during vulnerable periods, and enhance unattended infant safety by preventing potential undetected SIDS. This system improves the identification of important SIDS risk factors, overcomes practical limitations, and offers a solution for real-time unattended infant safety, potentially advancing research in the field of health.

2 Materials and Methods

2.1 Data Collection

The data collection process in this research used a Logitech C270 webcam. The consideration for choosing this webcam camera was because the Logitech C270 has a resolution of 720p, has image capture at 30 frames per second (FPS), and has a fixed focus feature. According to research [14], choosing a good webcam camera can be determined by camera's resolution; apart from that, the fixed focus feature is also an option because the image must not be blurry for detection using the deep learning method. Then, the Logitech C270 webcam is directed towards the infant's crib, which has been filled with a baby doll. The baby doll, which functions as a substitute for the infant because of the difficulty of taking photos of babies in the hospital, is actually placed in the crib. Various simulated scenarios of the infant's movements and positions are captured to produce high-resolution images. These images form a diverse dataset, essential for training and testing AI-integrated camera systems. This approach ensures ethical data collection, prioritizing infant safety and well-being while enabling the development of effective systems to improve infant safety monitoring. Figure 1 provides an example image that illustrates the potential implications of research data for creating safer environments for infants during their vulnerable developmental stages.

In Fig. 1(A), (D) represent that the infant is not in a dangerous position or in a position that can cause SIDS, while in (B), (C), and (E), the infant is in a dangerous position. In (B), the infant is said to be in a position that can cause SIDS because the infant's respiratory system can be disrupted in the prone position [15]. In (C), the infant's respiratory system could also be obstructed since it is covered by a blanket. While in (E), the falling risk is indicated due to the infant's spontaneous movement and poor balance. For the capturing image scenario, capturing is done clockwise around each corner; for variations, there are 3 variations, namely on the left side, right side, and middle side. For the total number of images taken, there were 443 sample images.

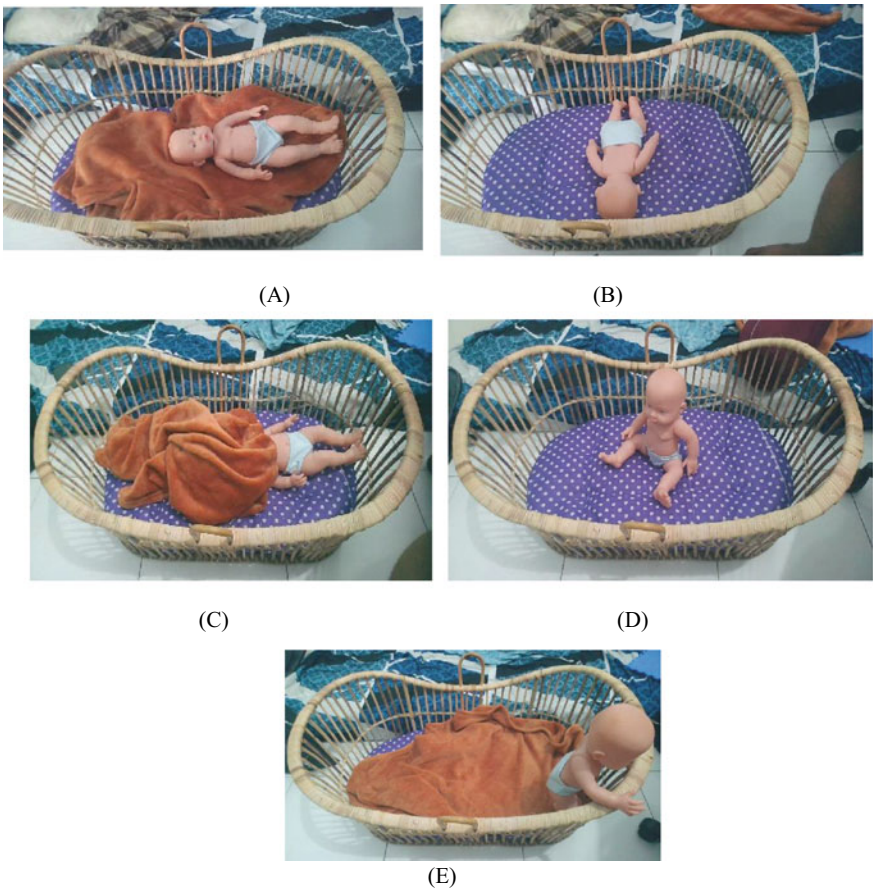


Fig. 1 Sample of infant's condition with respected label: **a** sleeping, **b** proning, **c** covered by a blanket **d**, sitting, **e** standing near a fence

2.2 Data Preprocessing

Upon acquiring the image data, the subsequent step involves performing image data labeling using a tool. This data labeling process entails drawing a bounding box line around the target object and aligning it accurately for precise labeling. Subsequently, a label name is assigned to the labeled object based on the captured image data. Notably, for this study, the label format employed is Pascal VOC, a format renowned for its emphasis on object detection, image classification, and segmentation tasks [16]. Following the labeling process, the subsequent stages encompass several additional processing steps, as illustrated in Fig. 2.



Fig. 2 Data preprocessing process

Once the data labeling process is accomplished, the dataset undergoes a crucial augmentation stage aimed at mitigating overfitting risks, given the inherent susceptibility of CNN deep learning models to this phenomenon [17]. The augmentation process entails a series of transformations designed to augment the dataset's diversity and enrich its content. One such technique involves resizing the images, a strategic approach to expand the scope of variations within the dataset. By introducing different image sizes, the augmented dataset provides the model with a more comprehensive and encompassing training experience, bolstering its ability to generalize effectively to novel and previously unseen data instances. Consequently, the model becomes more robust and less prone to overfitting.

The following step in the process entails the judicious separation of the dataset, whereby a strategic allocation of 80% is assigned to the training data, while the remaining 20% is purposefully designated for the validation data. This astute partitioning of the dataset serves a pivotal purpose, seeking to meticulously evaluate and scrutinize the model's performance and efficacy post-training [18]. By meticulously setting aside 371 samples for the training phase and allocating an additional 72 distinct samples for validation, this paper ensures a well-curated and diverse dataset that optimizes the model's capacity to generalize effectively and adeptly in real-world scenarios. This deliberate approach to data splitting empowers the researchers to robustly assess the model's ability to comprehend new and unseen patterns while also minimizing the risk of overfitting and enhancing the model's predictive capabilities.

Once all the data preprocessing procedures have been meticulously executed and the necessary data transformations and adjustments have been successfully applied, the dataset becomes aptly prepared for the commencement of the training phase, where cutting-edge deep learning models will be employed. This pivotal stage involves feeding the data into the neural network, enabling it to iteratively learn from the patterns, features, and underlying relationships within the dataset. Through the process of training, the model fine-tunes its internal parameters and optimizes its performance to achieve accurate predictions and classifications on unseen data points. The trained model then becomes a valuable tool for making informed decisions, recognizing patterns, and solving complex problems in various domains, ranging from computer vision and natural language processing to speech recognition and recommendation systems.

2.3 Deep Learning Model Selection

The selection of an appropriate deep learning model holds paramount significance, as it profoundly influences the entire implementation process. Among the various methodologies employed for image processing, CNN stands out as a popular and versatile choice. CNN encompasses a plethora of models, each catering to specific tasks and exhibiting distinct characteristics. Before finalizing the ideal model for the application, it is imperative to conduct meticulous benchmarking, rigorously evaluating several CNN models to gauge their performance and suitability. The benchmarking process entails the inclusion of models with manageable sizes, as larger models often incur extended processing times on the development board. Considering the criticality of real-time processing to prevent SIDS, minimizing the processing duration becomes a primary concern while simultaneously ensuring that the selected model attains commendable accuracy levels. Drawing insights from pertinent research endeavors [19].

The critical process of model selection, as evidenced by the comprehensive analysis detailed in Table 1, has led to the discerning choice of MobileNetv2 as the most fitting candidate. One of the key factors driving this decision is the model's impressively compact size, a mere 13 MB, significantly smaller than its counterparts like Inceptionv3, ResNet18, and ResNet50. While it is acknowledged that MobileNetv2's accuracy may not match that of ResNet50 and Inceptionv3, this drawback is mitigated by the realization that the former still exhibits commendable performance. This strategic trade-off between model size and accuracy is an essential consideration, especially given the constraints of real-time processing on the development board for SIDS detection. Furthermore, in the comparative analysis, MobileNetv2 emerges victorious in the accuracy contest when compared to ShuffleNet, despite the former's relatively larger size. This insightful revelation highlights the significance of prioritizing accuracy without unduly compromising computational efficiency, as this delicate balance holds the key to the model's real-world applicability and effectiveness.

Table 1 CNN model benchmarking

Network	Layers	Learnable layers	Network size (MB)	Parameters (Millions)	Accuracy (%)
Inceptionv3	315	48	89	23.9	92.90
ResNet18	71	18	44	11.7	90.00
ResNet50	117	50	96	25.6	92.42
MobileNetv2	154	53	13	3.5	90.81
ShuffleNet	172	50	5.4	1.4	86.06

Table 2 Development board benchmarking

	Raspberry Pi 4	Jetson Nano
Performance	13.5 GFLOPS	472 GFLOPS
CPU	Quad-core ARM CortexA72 64-bit @1.5 GHz	Quad-Core ARM Cortex-A57 64-bit @1.42 GHz
GPU	Broadcom Video Core VI (32-bit)	NVIDIA Maxwell w/128 CUDA cores @921 MHz
Memory	8 GB LPDDR4	4 GB LPDDR4 @1600 MHz, 25.6 GB/s
Price	$35	$89

2.4 Development Board Benchmarking

The process of selecting the most suitable development board for AI and machine learning endeavors demands a comprehensive and meticulous benchmark analysis, where critical factors such as compatibility, support for CNN models, computational performance, power efficiency, memory handling, and processing speed are carefully evaluated. In this study, a rigorous comparison was conducted between two widely utilized development boards, namely the Raspberry Pi 4 and the Jetson Nano (Table 2). This comparison aims to unravel their strengths, weaknesses, and overall capabilities. The insights derived from this comparative assessment provide invaluable guidance to developers and researchers, empowering them with the knowledge required to make well-informed decisions and effectively optimize AI applications for a diverse range of tasks and scenarios. By delving into the nuanced performance metrics of these development boards, advancements in AI technologies are fostered, paving the way for the seamless integration of powerful and efficient CNN models into real-world applications.

The benchmark analysis conducted in [20] reveals that the Jetson Nano outperforms the Raspberry Pi 4 in processing CNN deep learning models. The Jetson Nano demonstrates higher accuracy and significantly shorter processing times, owing to its powerful GPU architecture and support from the TensorRT engine. In contrast, the Raspberry Pi 4's limitations stem from the absence of a dedicated GPU and TensorRT support, making it less proficient in handling intensive CNN computations. Despite its higher cost, the Jetson Nano's superior performance justifies its value proposition for sophisticated applications where accuracy and processing speed are crucial.

2.5 Training Process

During this critical stage, the pre-processed data will undergo training utilizing the MobileNetv2 architecture. Leveraging the remarkable capabilities of the Jetson Nano, which is equipped with a Maxwell 128 CUDA Core GPU, the training process is poised to outperform that of the Raspberry Pi significantly. This advantage in

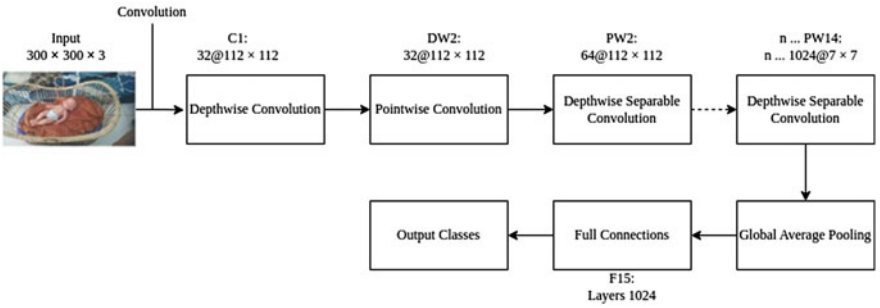


Fig. 3 Training on MobileNetv2 process

computational power and parallel processing capabilities contributes to faster and more efficient training, enabling the model to attain optimal performance levels. The intricacies of the training process, executed with the MobileNetv2 architecture, will be elucidated comprehensively in Fig. 3, offering valuable insights into the model's convergence and training dynamics.

In Fig. 3, the input image data with a size of 300 × 300 pixels undergo a series of transformations through three layers of the MobileNetV2 architecture. First, the image is divided into 112 × 112 pixels with a filtering value of 32 through the depth convolution method. Next, the image is processed through point convolution, retaining its size and filtering value. Then, deep separable convolutions are applied, maintaining the same size but with a different filtering value of 64. This process continues until the image size reaches 7 × 7 with a filtering value of 1024. After that, a global average merger is performed after the convolutional layers. Finally, a technique is used to convert 2D feature maps to 1D vectors. This operation effectively reduces the spatial dimension of the feature map to one value per feature channel. Consolidating global averages is critical for spatial dimensionality reduction, resulting in a compact representation of features, which is important for a variety of classification tasks.

The integration of ReLU6 activation in the MobileNetV2 model serves a vital role in enhancing its efficiency and speeding up processing time [21]. This feature becomes particularly crucial when considering the deployment of MobileNetV2 on resource-constrained development boards like the Jetson Nano. By employing ReLU6, the model benefits from reduced latency and accelerated computations, which are essential for real-time applications.

Another notable aspect of MobileNetV2 is its implementation of the linear bottleneck technique. This approach has been addressed in prior research [6, 22], highlighting the significance of inputs and outputs between model layers. Simultaneously, the inner layer serves as a crucial element in transforming low-level input concepts into high-level descriptors. Analogous to the residual connections in traditional CNNs, the shortcut connections in MobileNetV2 expedite training and enhance overall accuracy.

Table 3 Training parameters best fitting

Variable	Value
Epochs	50
Num_validation	2
Num_workers	2
Batch_size	8

During the training stage, the process was conducted on the Jetson Nano development board, carefully configuring the necessary settings to ensure compatibility with MobileNetV2. This involved the installation of specific libraries tailored to Jetson Nano's requirements. Subsequently, the training commenced with meticulously selected parameters, as outlined in Table 3. This parameter plays an important role in influencing the dynamics of learning and the convergence of the model towards optimal performance. Apart from that, the parameter values below are the most suitable because the number of epochs has been adjusted based on the amount of data, and trial error has been carried out four times with an epoch size of 10 to 50 epochs, considering the batch size for each training and the amount of validation data. From several experiments in searching for these parameters by looking at previous research [23], the best fit was obtained as shown in Table 3.

After training the model, the confusion matrix in Table 4, offers a comprehensive view of its performance. It provides valuable insights into true positive, false positive, true negative, and false negative values, enabling the calculation of important metrics like accuracy, precision, recall, and F1-score. This analysis helps identify areas of strength and areas for improvement, allowing us to optimize the model's accuracy and suitability for the task at hand.

To carry out the evaluation, several essential formulas are used to calculate the results of accuracy, recall, accuracy, F1-score, and misclassification rate denoted as Eqs. 1, 2, 3, 4, and 5 respectively.

Precision is a performance metric that quantifies the count of anticipated values, indicating their true nature. In the given context, *TP* represents true positivity, which occurs when the actual value and the predicted value are both positive. On the other hand, *FP* stands for false positive, which happens when the actual value is positive, but the predicted value is false.

Table 4 Confusion matrix table

	Prone	Sit	Covered_obstacle	Danger	Sleep
Prone	0.96	0.00	0.00	0.00	0.01
Sit	0.00	0.98	0.00	0.00	0.00
Covered_Obstacle	0.00	0.00	0.89	0.00	0.09
Danger	0.00	0.00	0.00	0.92	0.00
Sleep	0.00	0.00	0.12	0.00	0.99

$$precision = \frac{[TP]}{[TP + FP]} \tag{1}$$

To make it easier to remember the fraction of correctly detected positives, recall statistics are employed as a metric of an algorithm's capacity to identify all positive situations. *FN* refers to a false negative, which occurs when the true value is positive, but the predicted value is negative. In other words, it signifies instances where the model fails to identify positive cases correctly.

$$recall = \frac{[TP]}{[TP + FN]} \tag{2}$$

After obtaining the precision and recall values, the F1 score can be calculated. The F1 score represents the balanced average of weighted precision and recall.

$$F1score = \frac{\left[precision \times recall\right]}{\left[precision + recall\right]} \tag{3}$$

Subsequently, the accuracy of the trained model can be calculated.

$$accuracy = \frac{[TP]}{\left[Allsamples\right]} \tag{4}$$

Once the accuracy value is obtained, the misclassification rate can be calculated.

$$MisclassificationRate = 1 - accuracy \tag{5}$$

The results yielded from Eqs. 1 to5 serve as indicators to determine if the model requires any development. These metrics, including accuracy, precision, recall, F1-score, and misclassification rate, provide valuable insights into the model's strengths and weaknesses in classifying and predicting instances. Through a comprehensive analysis of these indicators, a clearer understanding of the model's proficiency can be gained, identifying areas where it excels and aspects that may require refinement.

2.6 Deployment on Development Board

Moreover, to gauge the model's practical utility in real-world scenarios, it will undergo testing using a USB Webcam camera connected to the Jetson Nano. This camera will be strategically positioned to have a clear view of the crib, ensuring accurate capture of the infant's position. The MobileNetV2 model will be executed on the Jetson Nano development board, with an emphasis on boosting processing speed through the utilization of the TensorRT engine, known for its substantial optimization advantages [24]. Leveraging TensorRT will not only enhance model efficiency

but also result in improved performance and reduced memory consumption. The practical implementation process, illustrated in Fig. 4, outlines the steps involved in using the model in a real-world context. To provide a succinct overview of Fig. 4, the system commences by activating the webcam camera. Subsequently, the Jetson Nano initializes the model through the program, and the webcam's feed appears on the display with bounding boxes for classification. It also presents information about the frames per second (FPS) and displays the classification of the infant's position. The system will stop if there is a command from Hotkeys. In real-world applications, this system offers flexibility for integration, including via a smartphone or a buzzer, and enables the prediction results to be conveniently viewed on a smartphone, facilitating monitoring from a considerable distance.

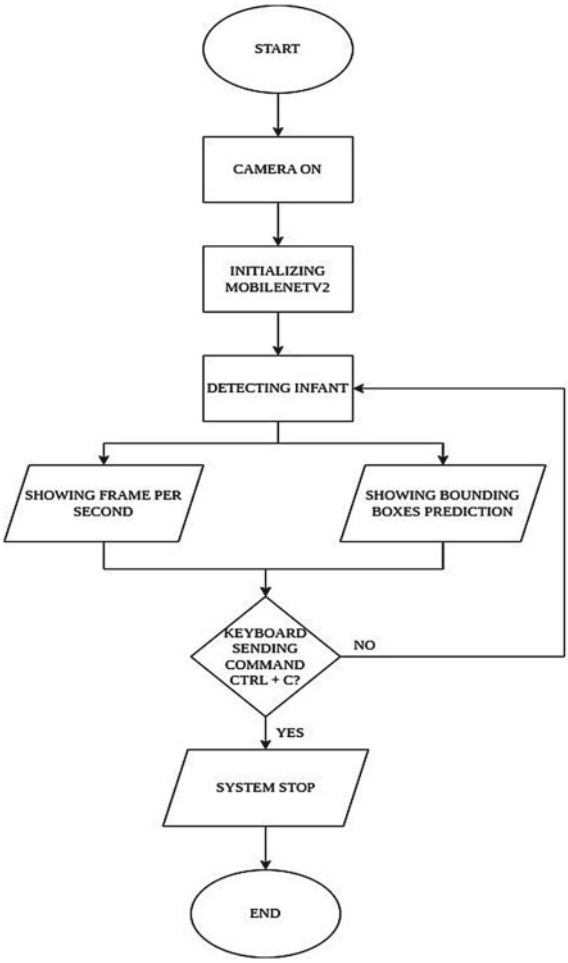


Fig. 4 System process

2.7 Software

For the labeling process, the software of choice is LabelImg, which enables the creation of accurate bounding boxes. Its compatibility with the Pascal VOC format streamlines the process, automatically generating files in the required format. Python 3.6.9 is selected as the programming language due to its pre-installed availability on the Jetson Nano development board and its extensive support for various deep learning frameworks. Among these frameworks, Pytorch is employed for creating deep learning models. Its user-friendly interface and abundant documentation make model development more accessible. Additionally, the vast and active community surrounding Pytorch ensures that any challenges encountered during the development process can be readily addressed, fostering a smooth and efficient workflow.

3 Result

The evaluation results obtained by analyzing the data from the confusion matrix using the formulas in Eqs. 1 to 5 provide a comprehensive assessment of the model's performance. These metrics enable us to determine whether the model can be deemed successful or requires further improvement.

The evaluation results extracted from Table 5 furnish compelling evidence substantiating the outstanding performance of the developed model. Notably, the model exhibits a remarkable average accuracy and F1-score of 94%, which signifies its adeptness in providing highly precise and accurate predictions across various object classes. This noteworthy level of accuracy attests to the model's robust architecture and meticulous training process. Moreover, let's delve into the specifics of the results for each class. For the "Class Name," precision scores are indicative of how well the model correctly identified instances of that class, with "Recall" denoting the proportion of actual instances within that class that the model successfully detected. The "F1-Score" represents a balance between precision and recall, providing a harmonized measure of the model's performance for each class. In this context, the model demonstrates exceptional precision, recall, and F1-score across all classes, illustrating its proficiency in distinguishing and accurately identifying objects. This comprehensive evaluation underscores the model's reliability and its capacity to minimize false predictions while effectively detecting objects within each class.

Furthermore, the outcomes showcased in Table 6 serve as a representative illustration of the model's performance when executed on the Jetson Nano development board. This table offers valuable insights into the model's ability to conduct real-time detection directly through the camera feed. It provides a breakdown of the confidence scores acquired for each class, where confidence scores exceeding 90% signify the model's proficient classification of the infant's position, accurately tracking the infant's movements with the bounding box. The overall average confidence score,

Table 5 Model evaluation result

Class name	Precision	Recall	F1-score
Prone	0.98	0.98	0.98
Sit	0.97	1.00	0.98
Covered_obstacle	0.90	0.86	0.88
Danger	1.00	1.00	1.00
Sleep	0.89	0.90	0.90
Avg. accuracy	0.94		
Misclassification rate	0.06		

Table 6 Real testing results

Class name	Avg. Conf. Score	Avg. FPS
Prone	0.99	36
Sit	0.99	32
Covered_obstacle	0.95	28
Danger	0.98	30
Sleep	0.97	31

computed across all classes, stands at an impressive 97.6%, underscoring the model's exceptional capability to comprehensively classify infants. Moreover, each model attains an FPS value surpassing 20, signifying smooth image processing operations. Although the "obstacle-covered" label yields a slightly lower average score compared to other labels, this can be attributed to the label's dual-object detection requirement, involving both infants and obstacles. Nevertheless, the average score of 31.4 across all labels demonstrates that image processing remains consistently efficient.

In contrast, when examining previous studies in the image processing domain [25–29], the YOLO framework emerges as the most widely adopted framework. Interestingly, even in the most recent research related to infant monitoring through image processing [30], the YOLO framework remains the preferred choice. This suggests that the YOLO framework holds substantial significance due to its prowess in three pivotal aspects: accuracy, speed, and predictive capabilities. While these findings from the previous study can serve as valuable reference points for this research, it's important to note that the aforementioned study did not provide specific details regarding the device employed or the speed of the image processing. To address this gap and offer a more comprehensive comparison, insights are also drawn from another study [31] that utilizes the YOLO framework. Although this research diverges from the healthcare domain, it delves into the performance attributes of the YOLO framework, thus contributing to a more comprehensive understanding of its capabilities beyond the specific health-focused context.

Table 7 (b) compiles the outcomes of prior research employing the YOLOv5 model, spanning from the lowest version (YOLOv5s) to the highest (YOLOv5x).

Table 7 YOLOv5 model performance [31]: YOLOv5 real testing FPS with Jetson Nano (a), YOLOv5 real testing confidence score with Jetson Nano (b)

(a)	
Model [31]	FPS
YOLOv5s	4.64
YOLOv5-mv3s	9.8
YOLOv5-mv3s (enhanced)	11.5
(b)	
Model [31]	Conf. Score
YOLOv5s	0.89
YOLOv5m	0.94
YOLOv51	0.92
YOLOV5x	0.94

Table 8 T-Test Result between [31] with MobileNetv2

Comparison list	t	α (%)	p-value	d
Model comparison	3.53	5	0.03	0.49
FPS comparison	5.39	5	0.03	3.12

Meanwhile, Table 7 (a) encapsulates previous research outcomes involving the optimization of the YOLOv5S model for the Jetson Nano development board. These outcomes establish a benchmark for contrasting the performance of the MobileNetv2 model employed in the current study, which was executed on the same development board. A paired t-test statistical analysis was conducted to assess the study's viability, allowing for an investigation into whether there exists a significant disparity in the performance of MobileNetv2 and YOLOv5 on the Jetson Nano platform.

The t-tests presented in Table 8 offer essential insights into the performance comparison between the MobileNetv2 and YOLOv5 models. These tests assess the significance of differences in critical metrics such as FPS and confidence scores. The t-statistic (T) quantifies these differences, while the significance level (alpha) sets the threshold for considering these differences statistically significant. The p-value, a key indicator, reflects the probability of observing such differences by random chance. A small p-value, in this case, indicates that observed disparities are unlikely to be mere chance occurrences. Moreover, the effect size (d) measures the practical significance of these differences.

4 Discussion

The training results achieved with the MobileNetv5 model demonstrate an average accuracy of 94% and a misclassification rate of 6%. This indicates that, overall, the MobileNetv2 model excels in making accurate predictions across all labels. The

results in Table 5 further support these findings, with precision values exceeding 85% for each label. This signifies a strong ratio of true positive predictions to the overall positive predictions for each label. Additionally, the recall values for each label are also above 85%, indicating high detection sensitivity for each label. Consequently, when there is a change in the infant's position, the MobileNetv2 model swiftly recognizes it. These two parameters are further validated by the F1-Score, which consistently exceeds 90%. This underscores the MobileNetv2 model's excellence in achieving both precision and recall in classification. Subsequently, upon implementation on the development board illustrated in Table 6, the mean confidence score for each label surpasses 90%. This signifies that the real-world implementation of the model with a webcam leads to effective classification, with the bounding boxes accurately tracking the object's movements. Simultaneously, the average FPS value for each label exceeds 20 FPS, underscoring that a higher FPS corresponds to a swifter image processing workflow.

Based on the results of the T-test presented in Table 8, we can reject the null hypothesis because the p-values are below the significance level (α). This signifies a substantial difference in performance between the MobileNetv2 and YOLOv5 models. The effect size (d) values further emphasize this difference. Specifically, the effect size for FPS is 3.12, indicating a significant positive impact of the MobileNetv2 model on performance compared to YOLOv5 in terms of processing speed. Conversely, the effect size for the confidence score is 0.49, suggesting a more moderate effect on both models in terms of object classification accuracy. Furthermore, the comparison between MobileNetv2 and YOLOv5 extends to memory and storage usage. MobileNetv2 boasts a compact network size of 13 MB, while even the simplest architecture of YOLOv5 demands a slightly larger 14 MB network size. Additionally, MobileNetv2 demonstrates efficiency in memory utilization.

However, this study has limitations. The MobileNetv2 model achieves a high FPS value, surpassing 20 FPS, thanks to the assistance of TensorRT. This high FPS rate enables real-time monitoring but is currently limited to specific development boards like the Jetson Nano. Implementing the MobileNetv2 model to achieve 20 FPS on TinyML devices is relatively uncommon at present. Some research has applied the MobileNet model to the Raspberry Pi, but the results still fall below 10 FPS [32]. TinyML solutions could potentially address this issue as they are compact, cost-effective devices, unlike the Jetson Nano.

Addressing the ethical dimensions of infant safety in AI monitoring is paramount. It involves careful consideration of potential false positives and negatives, emphasizing the need for AI systems to minimize errors that can cause undue distress or miss genuine threats. Equally important is discussing the consequences of relying solely on AI for infant monitoring, highlighting that it should complement, not replace, parental vigilance. Promoting the responsible use of AI by educating parents about its capabilities and limitations while reinforcing the crucial role of parental care is essential. These discussions must guide the ethical development and deployment of AI technology in infant safety, ensuring that it augments safety without diminishing the vital role of attentive parenting.

5 Conclusion

The primary objective of this study is to provide parents with an effective means of monitoring their infants. Parenting entails numerous responsibilities, and while tending to these, it's imperative to ensure the well-being of infants, who are particularly vulnerable due to their delicate bone structure and limited motor control. This research serves a crucial role in averting inappropriate sleeping positions and safeguarding infants from potential respiratory obstructions, thus mitigating the risk of sudden death. The methodology employed here leverages deep learning techniques for infant position detection, with the MobileNetv2 model being the chosen framework due to its compact network size, weighing in at just 13 MB, making it a more lightweight option compared to the commonly used YOLOv5 model for image processing. Consequently, the p-value results indicate improvement, particularly in the FPS metrics, where the MobileNetv2 model excels in image processing when compared to YOLOv5. A high FPS value signifies rapid image processing, resulting in smoother video recordings and quicker decision-making by the MobileNetv2 model. This speed is crucial in responding promptly to potentially hazardous situations, ensuring parents can take swift preventive measures. Looking ahead, there is optimism that high FPS image processing can be achieved on TinyML devices in the future, although current limitations hinder rapid image processing in TinyML. It's plausible that advancements in chip technology or the development of dedicated hardware, such as a specialized VGA, may facilitate the integration of high-FPS image processing into TinyML devices.

References

1. Yasumitsu-Lovell K et al (2021) Birth month and infant gross motor development: results from the Japan environment and Children's study (JECS). PLoS One 16(5):e0251581. https://doi.org/10.1371/journal.pone.0251581
2. Suhartini B (2017) Stages of infant motor development. Yogyakarta State University Journal, Indonesia
3. Abdullah K et al (2022) The concept of the middle body (Midline) in development baby motoric. Muhammadiyah Surabaya University Journal, Indonesia
4. Park S et al (2022) The global burden of sudden infant death syndrome from 1990 to 2019: a systematic analysis from the global burden of disease study 2019. QJM 115(11):735–744. https://doi.org/10.1093/qjmed/hcac093
5. Jullien S, Huss G, Weigel R (2021) Supporting recommendations for childhood preventive interventions for primary health care: elaboration of evidence synthesis and lessons learnt. BMC Pediatrics 21. BioMed Central Ltd. https://doi.org/10.1186/s12887-021-02638-8
6. Zayed Y, Hasasneh A, Tadj C (2023) Infant cry signal diagnostic system using deep learning and fused features. Diagnostics 13(12):2107. https://doi.org/10.3390/diagnostics13122107
7. Talukdar S, Saha S (2021) Intelligent baby monitoring system using blynk, 51–64. https://doi.org/10.1007/978-981-15-8752-8_6
8. Visvesvaran C et al (2021) IoT based smart baby monitoring. Second Int Conf Electron Sustain Commun Syst (ICESC) 2021:1–6. https://doi.org/10.1109/ICESC51422.2021.9532997

9. Durga S, Itnal S, Soujanya K, Basha CZ, Saxena C (2021) Advanced and effective baby care monitoring Smart cradle system using Internet of Things. In: 2021 2nd international conference on smart electronics and communication (ICOSEC), pp 35–42. https://doi.org/10.1109/ICOSEC51865.2021.9591955
10. Symon A, Hassan N, Rashid H, Ahmed I, Reza S (2017) Design and development of a smart baby monitoring system based on raspberry Pi and Pi camera. https://doi.org/10.1109/ICAEE.2017.8255338
11. Kong S, Zhou C, Sun J (2023) A review of the application of convolutional neural networks in object detection. In: Q, Wang W, Mu J, Liu X, Na Z (eds), Artificial intelligence in China, Liang, Singapore: Springer Nature Singapore, pp 177–186
12. Bharati V (2021) An efficient edge deep learning computer vision system to prevent sudden infant death syndrome. IEEE Int Conf Smart Comput (SMARTCOMP) 2021:286–291. https://doi.org/10.1109/SMARTCOMP52413.2021.00061
13. Alam H et al (2023) IoT based smart baby monitoring system with emotion recognition using machine learning. Wirel Commun Mob Comput 2023:1175450. https://doi.org/10.1155/2023/1175450
14. Puppim de Oliveira D, Pereira Neves dos Reis W, Morandin Junior O (2019) A qualitative analysis of a USB camera for AGV control. Sensors 19(19):4111. https://doi.org/10.3390/s19194111
15. Priyadarshi M, Balachander B, Sankar MJ (2022) Effect of sleep position in term healthy newborns on sudden infant death syndrome and other infant outcomes: a systematic review. J Glob Health 12. https://doi.org/10.7189/JOGH.12.12001
16. Everingham M, Van Gool L, Williams CKI, Winn J, Zisserman A (2010) The pascal visual object classes (VOC) challenge. Int J Comput Vis 88(2):303–338. https://doi.org/10.1007/s11263-009-0275-4
17. Alomar K, Aysel HI, Cai X (2023) Data augmentation in classification and segmentation: a survey and new strategies. J Imaging 9(2):46. https://doi.org/10.3390/jimaging9020046
18. Xu Y, Goodacre R (2018) On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. J Anal Test 2(3):249–262. https://doi.org/10.1007/s41664-018-0068-2
19. Alzahrani S, Al-Bander B, Al-Nuaimy W (2021) A comprehensive evaluation and benchmarking of convolutional neural networks for melanoma diagnosis. Cancers 13(17):4494. https://doi.org/10.3390/cancers13174494
20. Süzen A, Duman B, Şen EB (2020) Benchmark analysis of Jetson TX2, Jetson nano and raspberry PI using deep CNN.https://doi.org/10.1109/HORA49412.2020.9152915
21. Dang TV (2022) Smart home management system with face recognition based on ArcFace Model in deep convolutional neural network. J Robot Control (JRC) 3(6):754–761. https://doi.org/10.18196/jrc.v3i6.15978
22. Nyoman P, Negara PK (2021) Deteksi Masker Pencegahan Covid19 menggunakan convolutional neural network berbasis android. J RESTI (Rekayasa Sist Teknol Inf) 5(3):576–583
23. Afaq S, Rao S (2020) Significance of Epochs on training a neural network. Int J Scientific Technol Res 9:485–488. https://api.semanticscholar.org/CorpusID:225647672
24. Rahmaniar W, Hernawan A (2021) Real-time human detection using deep learning on embedded platforms: a review. J Robot Control (JRC) 2(6). Department of Agribusiness, Universitas Muhammadiyah Yogyakarta, pp. 462–468Y. https://doi.org/10.18196/jrc.26123
25. Tian M, Liao Z (2021) Research on flower image classification method based on YOLOv5. J Phys Conf Ser 2024(1):012022. https://doi.org/10.1088/1742-6596/2024/1/012022
26. Mohiyuddin A et al (2022) Breast tumor detection and classification in mammogram images using modified YOLOv5 network. Comput Math Methods Med 2022:1359019. https://doi.org/10.1155/2022/1359019
27. Yang G et al (2020) Face mask recognition system with YOLOV5 based on image recognition. In: 2020 IEEE 6th international conference on computer and communications (ICCC), pp 1398–1404. https://doi.org/10.1109/ICCC51575.2020.9345042

28. Majeed F et al (2022) Investigating the efficiency of deep learning-based security system in a real-time environment using YOLOv5. Sustain Energy Technol Assess 53:102603. https://doi.org/10.1016/j.seta.2022.102603
29. Wu B, Pang C, Zeng X, Hu X (2022) ME-YOLO: improved YOLOv5 for detecting medical personal protective equipment. Appl Sci 12(23):11978. https://doi.org/10.3390/app122311978. (sudahIEEE)
30. Sanjay NS, Ahmadinia A (2019) MobileNet-Tiny: a deep neural network-based real-time object detection for rasberry Pi. In: 2019 18th IEEE international conference on machine learning and applications (ICMLA), pp 647–652. https://doi.org/10.1109/ICMLA.2019.00118
31. Zhao X, Huang Z, Ye L, Lv Y (2022) Real-Time detection method for submarine pipeline leakage based on deep learning and Jetson nano in OCEANS. Hampton Roads 2022:1–10. https://doi.org/10.1109/OCEANS47191.2022.9977048
32. Chung GC, Teong KV,. Tan SF, Sundarajoo RA (2022) IoT-based autonomous baby monitoring system using mobile app control. In: Alfred R, Lim Y (Eds), Proceedings of the 8th international conference on computational science and technology. Singapore: Springer Singapore, pp 183–193

Visualization of Data from a Multifunctional Surgical Device for Measuring Forces and Torques Using the Violin Diagram Method

Mikhail A. Solovyev, Andrey A. Vorotnikov, Andrey A. Grin, Yuri V. Poduraev, Anton Y. Kordonsky, and Oleg V. Levchenko

Abstract Recently, there has been a trend in the field of surgery towards the development of smart medical devices. Sensors of various functionalities are embedded in medical-surgical devices, which makes it possible to read data during surgical interventions and, based on the data obtained, track critical situations and analyze the actions of the surgeon. One of the recent trends is the introduction of multicomponent force-torque sensors in surgical instruments. However, when processing data from such sensors, a problem arises with the visualization of data in a form convenient for human perception. The purpose of this article is to describe an approach to presenting data in the form of violin diagrams for analyzing the data of forces and moments occurring on medical instruments with built-in multi-component force sensors. The main criteria necessary for obtaining the maximum information content of visualization are highlighted, based on which the synthesis of the visualization method is carried out. The proposed method is also tested for various surgical effects on patient tissues. In particular, the results of testing the visualization method are described by experimenting using a multifunctional device for measuring forces and moments on back phantoms, followed by data visualization. The possibility of using the data visualization method for various methods of surgical interventions, as well as the possibility of comparing numerical parameters, is evaluated.

M. A. Solovyev (✉) · A. A. Vorotnikov · A. A. Grin · Y. V. Poduraev · O. V. Levchenko
Federal State Budgetary Educational Institution of Higher Education "Russian University of Medicine" of the Ministry of Health, Moscow, Russia
e-mail: mikke2@yandex.ru

A. A. Grin · A. Y. Kordonsky
Sklifosovsky Research Institute of Emergency Medicine, Moscow, Russia

M. A. Solovyev · A. A. Vorotnikov · Y. V. Poduraev
Moscow State University of Technology "STANKIN", Moscow, Russia

A. A. Grin
Moscow Healthcare Department, Moscow, Russia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_20

Keywords Visualization · Data processing · Force measurement · Surgical instrument · Graph

1 Introduction

Modern trends in surgery are aimed at reducing trauma and moving towards more precise and minimally invasive treatment of patients. Similar results are achieved through the introduction of more advanced surgical instruments and sophisticated robotic systems. The dynamics of the development of devices of this type in various areas of surgery is relatively high, and there is a high need to find and develop better solutions [1, 2]. One such solution is the transition from conventional surgical instruments to "smart" surgical instruments. This type of tool retains the functionality of the previous generation while adding the ability to record and transmit data from tools through various built-in sensors. Thus, it becomes possible to obtain additional information about the progress of the operation. Such devices make it possible to implement new possibilities, such as monitoring the correctness of the surgical procedure [3], and training of medical personnel [4]. A promising direction in this area is the integration of multicomponent force-torque sensors into surgical instruments to track the forces applied to the tissues of patients. This approach makes it possible to collect data both during operations with normal tissues [5] and with inert tissues [6] of the patient. The data obtained are used to refine the biomechanical models of human tissues [7] and make it possible to more accurately determine the initial design parameters for more advanced medical systems [8–11]. However, the use of force-torque multicomponent sensors creates the problem of data visualization. The presence of a large number of components of the sensor and the data obtained in the form of a set of noisy analog signals of different amplitude and time complicate the task of finding a universal, informative method of data presentation.

There are a large number of approaches to the visualization of multiparameter data [12, 13], however, for each area, it is necessary to search for the optimal method depending on the required criteria. A number of works have already considered the problem of visualization of force effects obtained with the help of surgical instruments. For example, in [14], a system with strain gauges on surgical forceps is presented. However, in this study, only data in the form of a single analog signal is considered, and the data is presented in the form of a table of correspondence of one analog voltage to force. The papers [15, 16] present a description of the measurements of the forces that arise on the scalpel when cutting tissues. However, in this case, researchers were mainly interested in one axis along which the force was directed, so this method is not suitable for multicomponent sensors. Studies [17, 18] present visualization complexes for surgical instruments, however, the authors in these articles focus on the three-dimensional representation of instruments. The study [19] describes a study with a surgical device with feedback on a multicomponent force sensor, however, the data themselves are presented in the form of graphs, which can

complicate the comparison of data in the case of noisy graphs or different values taken over time.

Thus, in order to compare data from multicomponent force-torque sensors for different durations of force actions by one type of tool, as well as for data from different types of tools, it was necessary to find a universal representation of the output data. The authors of this article propose a solution to this problem by applying the methods of statistical analysis and visualization methods to obtain informative diagrams that allow comparison of force effects. Within the framework of this article, the main required criteria for choosing an imaging method will be highlighted, based on the criteria, an imaging method will be synthesized and tested using a surgical device with a multicomponent force-torque sensor. An important result obtained from the synthesis of the visualization method is its ability to be used for comparing different types of surgical interventions, intuitive clarity, and the ability to integrate the results of the method into various software environments.

2 Materials and Methods

The synthesis of the visualization method will be carried out based on the data of a multifunctional device for measuring forces and moments in spinal neurosurgery [20]. Its appearance is shown in Fig. 1 This device is a set of interchangeable surgical instruments that are mated with the handle through a mechanical connection with a compact 6-component force-torque sensor. The authors of this article previously made a prototype device with a data processing system and tested it, which proved its full performance. The resulting device is capable of measuring forces and moments along three Cartesian axes with a maximum force per axis of 200 [N] and a maximum torque of 5 [Nm] to 20 [Nm], depending on the selected surgical tool.

The primary task of finding a new visualization method was to highlight the key characteristics that needed to be reflected in a graphical representation with



Fig. 1 Appearance of a multifunctional device for measuring forces and moments in spinal neurosurgery

maximum information content. In the case of processing data on forces and torques from a multi-component sensor, the following were chosen as critical characteristics:

- The maximum obtained value of the magnitude of forces and torques
- Range of measured forces within one dimension
- Characteristics of the Measurement Obtained

The first two characteristics will allow you to compare the data with each other in absolute numbers. Thus, when considering a large number of similar measurements, it becomes possible to compare them and determine the correspondences between efforts quickly. However, in the case of presenting characteristics data for all components obtained from the sensor, the final image will look overloaded. To improve the display method, it was necessary to adjust the values according to some transformation functions to make them more convenient for comparison, namely, to resort to the data normalization operation [21]. The goal of this normalization was to minimize the information components from the input data. For this, it was necessary to consider the general normalization problem for a multifunctional surgical device for measuring forces and torques. Data acquisition in the device was carried out using a built-in six-component force-torque sensor. At the output of the data processing system, the values from each sensor component were collected, which can be represented as a six-component vector $\overrightarrow{F_i}$. This vector can be defined as:

$$\overrightarrow{F_i} = \left(F_{xi}, F_{yi}, F_{zi}, M_{xi}, M_{yi}, M_{zi}\right)^T \tag{1}$$

where: Fxi, Fyi, Fzi are the components of the i-scale force vector in the Cartesian coordinate system, Mxi, Myi, Mzi are the components of the i-scale torques vector in the Cartesian coordinate system. To reduce the number of information components, the display of vector values was presented in their modules.

$$\begin{cases} F_{i\sum} = \sqrt{F_{xi}^2 + F_{yi}^2 + F_{zi}^2} \\ M_{i\sum} = \sqrt{M_{xi}^2 + M_{yi}^2 + M_{zi}^2} \end{cases} \tag{2}$$

The transition to the modules of vectors allows not only reduces the number of components from six to two but also retains information content since the module of the vector will reflect the magnitude of the force vector applied by the surgeon to the instrument. By highlighting the range of moduli of the obtained vectors on the graphs, it becomes possible to compare the applied total force and torque for different stages of the surgical intervention [22] Thus, two key characteristics of the force effect are reflected: the maximum value obtained and the range of the measured forces. However, the third characteristic, associated with the peculiarity of the obtained values, remains not displayed. This characteristic should provide information about dynamic changes in values. So, for example, in the case of considering two force actions in which the maximum applied force is the same, but the time of application of these efforts is different, information about the maximum efforts and ranges will not

give an understanding of the difference between the effects. For this, it was necessary to find an additional method capable of reflecting the dynamics of changes.

The limited visual presentation of data complicates the search for this method. Since the reflection of the ranges of normalized parameters was chosen as the basis for the visualization method, the method reflecting the dynamics of changes should be reflected next to them and be either two-dimensional or three-dimensional. Additional complexity is introduced by the possibility of a significant difference in the amplitudes of the measurable quantities. Suppose the method is presented in the form of a graph of the dependence of the modulus of the force vector on time. In that case, if the measured values differ by orders of magnitude, it will be impossible to reflect the graphs for convenient comparison correctly. In this case, the graphs themselves can be significantly noisy, complicating the comparison process. Thus, it was possible to distinguish the following requirements for the visual display of the dynamic component of the force measurement:

- Ability to embed in classical methods of visual display in 2D and 3D spaces
- Independence from the amplitude of values
- Independence from signal noise

This solution was found by presenting the data in the form of violin diagrams. These charts combine two types of graphical presentation of information: a bar chart that reflects the maximum and minimum value of the measured value and a histogram of the distribution of values within a given range of values. Column charts are presented as vertical bars, and distribution histograms are symmetrical graphs on the sides of these bars. The amplitude of the graph corresponds to the frequency of occurrence of a given value. The use of histograms, in this case, is due to getting rid of data noise and the ability to represent values with amplitudes that differ by order of magnitude or more on a single-dimensional scale. Also, such a representation allows you to visually highlight the characteristic force ranges associated with prolonged force effects on a surgical instrument as areas with an increased concentration of values around the average value of this value. Thus, when examining one violin diagram, it becomes possible to obtain a significant amount of information about the nature of the force effect. Figure 2 shows an illustrative example of the use of this visualization method. Graphs showing the dependence of the components of the force vectors and the torque on time are presented on the left. Due to the high variance of the data, it is difficult to judge the nature of the force effect and to highlight useful information. In addition, to get a complete picture of the nature of the force, it is necessary to reduce the vector components to the modulus of this vector. Suppose the value of the module is reflected as another graph. In that case, this will lead to an even greater complication of the visual component and complexity in working with data in general. A completely different situation arises in the case of data presentation in the form of violin diagrams. The graphs to the right show the representation of the same data in the form of violin diagrams. We can observe the range of values for the vector of forces and values in modulus, from the minimum obtained value to the maximum. In this case, the histogram of the distribution of values indicates two

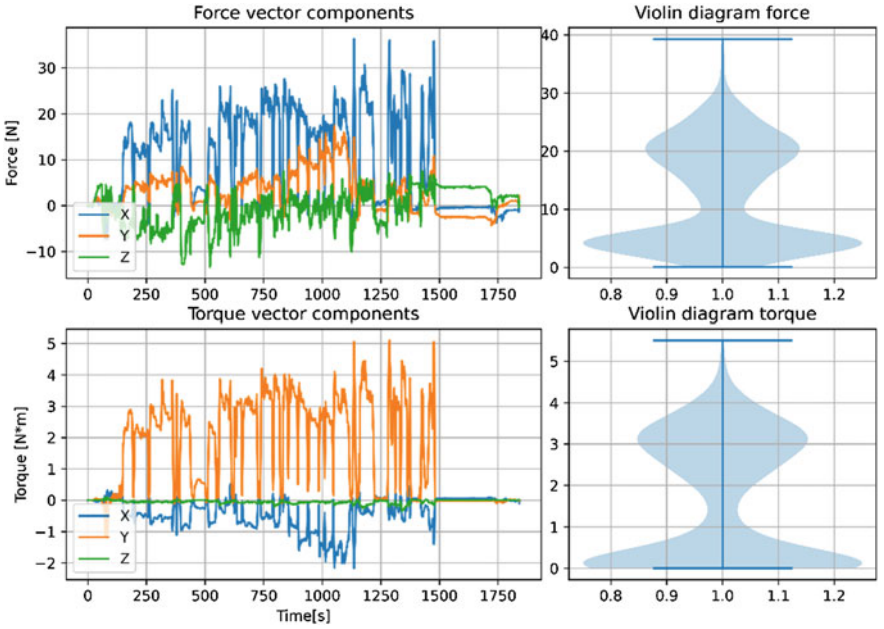


Fig. 2 View of the graphs in their original form and the form of a violin diagram

different influences, where the force was close to zero and a real force effect with the average value obtained.

Thus, it can be noted that the presentation of data in the form of violin diagrams allows you to reflect all the essential characteristics of the force impact. In this case, the resulting force patterns are compact and conveniently arranged on a standard graph.

3 Results

3.1 Testing the Method on Different Types of Surgical Interventions

To test the efficiency of this method of data presentation in practice, three experiments were carried out using a multifunctional device for measuring forces and torques on surgical instruments. Instruments for carrying out transpedicular fixation of the injured vertebra were chosen as instruments. The first experiment was to prove the possibility of determining the differences in the dynamic component on violin force action diagrams for various types of surgical actions. For this purpose, two forces were applied to the silicone phantom of the back with a spine model using a

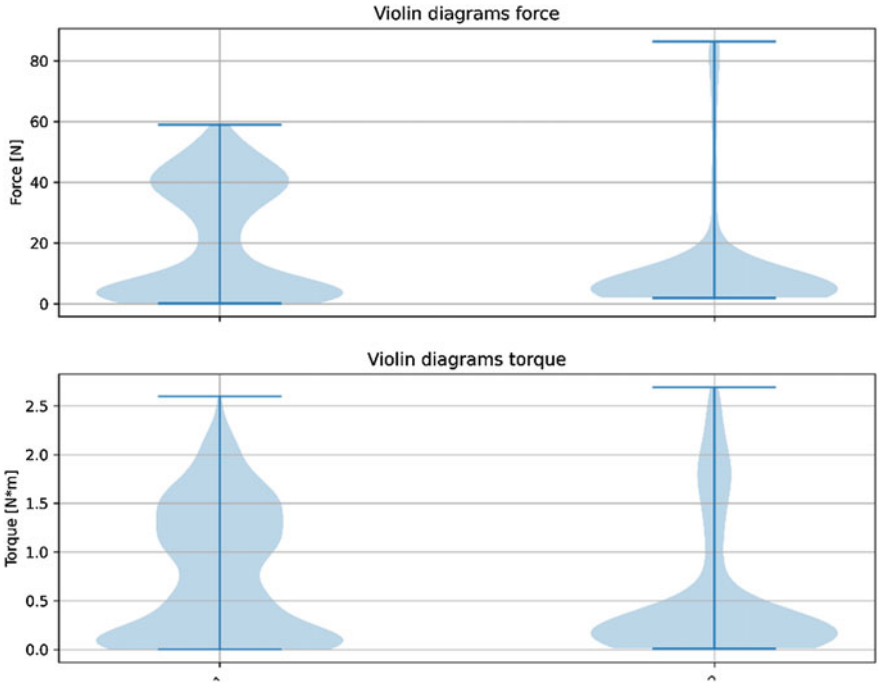


Fig. 3 Force data in the form of violin diagrams for different surgical interventions

multifunctional device. Spine Holder Set, L1-Sacrum from sawbones was used as a phantom with a silicone layer to simulate back muscles. During the first forceful action, an L-shaped retractor was used as a surgical instrument to simulate muscle abduction on a silicone back phantom. During the second exposure, an awl-type instrument was used to simulate the piercing of the bone tissue of the spine. The results of the experiment are shown in Fig. 3.

We can observe significant differences in the histograms of force patterns of actions both in terms of forces and in torques. This is due to the nature of the required force from the surgeon. In the case of pulling back soft tissues (Fig. 3 left), it is necessary to act on the instrument for a long time. Thus, an area with a high density of values close to the maximum value of the force appears on the violin diagram. In the case of the violin diagram connected by the imitation of the puncture of the bone tissue (Fig. 3 right) with an awl, a low density of values is observed at the upper values of the force. This is due to the fact that the puncture procedure requires a short duration of force on the vertebral tissue. However, both force patterns retain an increased density of values for the lower values of the force associated with tool movement. Thus, it was confirmed that it is possible to determine the differences in the dynamic component of force effects for various surgical interventions.

3.2 Testing the Method for Effective Data Comparison

The second experiment was conducted to test the possibility of effective data comparison in the framework of force effects with one type of surgical instrument. For this purpose, the imitation of muscle abduction was carried out several times on a silicone phantom of the back. The Spine Holder Set, L1-Sacrum from sawbones was also used as a phantom with a silicone layer to simulate the muscles of the back. There were three impacts in total. The result obtained is shown in Fig. 4. They were carried out with an L-shaped retractor. The graphs show how with an increase in the maximum applied force at each stage, the maximum torque increases. At the same time, on the histograms of the distribution of the force of both instruments, there are two areas with a high density of values. The first area, which is close to the minimum of the measured force, corresponds to the values in which the surgeon brought the instrument to the tissues. During these periods of time, the instrument did not come into contact with the phantom tissues, and being without a fulcrum was subjected to the effects of gravity. For the L-shaped retractor, these effects are seen as an increase in the distribution density in the range from 0 to 15 [N]. A similar picture is observed for torques. If the tools have a fulcrum, this effect is significantly reduced. During the withdrawal of the back muscles, the surgeon applies a long-term force, which is expressed in the form of a distribution of a large number of forces close to the maximum values. The large scatter of these values is due to the nature of the impact, which requires the surgeon to change the strength and direction of the force. The maximum value of the reduced force vector for this type of impact was determined to be 85.4 [N], and the maximum value of the torque was 5.2 [Nm].

3.3 Testing the Possibility of Displaying Changes in the Mechanical Properties of Patient Tissues

The last experiment was to test the possibility of displaying, using violin diagrams, changes in the mechanical properties of the patient's tissues when they are exposed to surgical instruments. For this purpose, an imitation of the pedicle screw screwing into the phantoms of the vertebrae of the back was carried out. As phantoms of the vertebrae, polyurethane foam bars with variable mechanical characteristics were used, obtained by using different percentages of components in the material mixture. Depending on the percentage of the sum of two components of molded polyurethane foam A2 + B2 from the mass of the entire mixture of material, it was possible to significantly change the density of the final vertebral phantom. For the experiment, the following values of the percentage of the sum of the mixture of components were chosen: 25%, 35%, 40%, 45%, 50%. The higher the percentage in the final material, the lower the density of the resulting phantom material. The experiment itself was carried out by manually screwing the pedicle screw to the same depth using

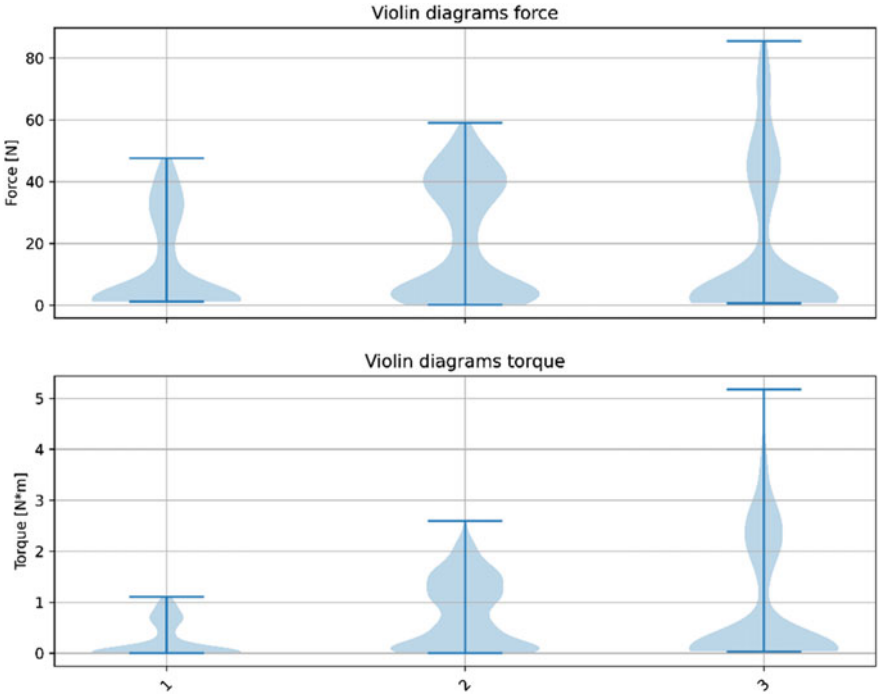


Fig. 4 Force pictures in the form of violin diagrams for several influences with one surgical instrument

a multifunctional surgical device. The result of the obtained data in the form of violin diagrams is shown in Fig. 5.

The resulting violin diagrams show a linear increase in the maximum applied values of forces and moments with a decrease in the percentage of the sum of the mixture of components. This shows an increase in the maximum values from 68.3 [N] for forces, 1.93 [Nm] for moments 158.2 [N] for forces, and 4.1 [Nm] for moments. Thus, the ability to display changes in the mechanical properties of patient tissues using violin diagrams was also confirmed.

4 Discussion

Thus, by processing the data of the experiments described above, the possibility of an effective comparison of the force effects of a surgical instrument with sensory elements has been shown. The resulting display method meets the requirements for the implementation of comparison when embedded in classical visual display methods, independence from the amplitude of values, and noisiness. Moreover, an essential feature of this method is the possibility of distinguishing different methods

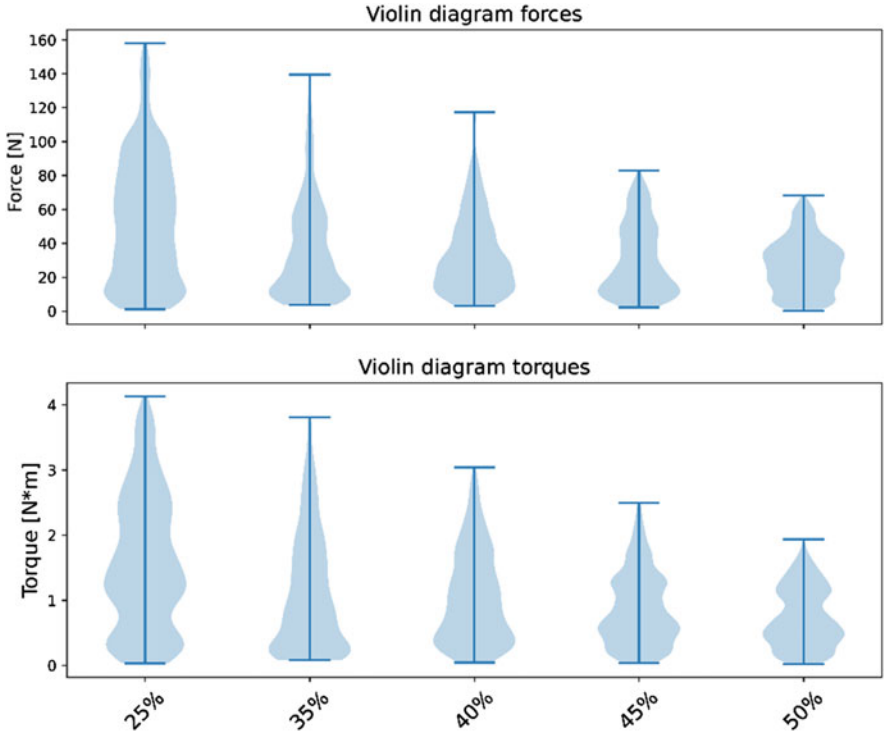


Fig. 5 Data in the form of violin diagrams of an experiment on screwing a pedicle screw into vertebral phantoms with different mechanical properties

of surgical intervention by types of histograms of the distribution of values. When looking at similar studies, one can see a trend related to assessing the maximum characteristics of the mechanical properties of fabrics and materials. Thus, in [23], a comparison is made of the properties of a created vertebra phantom and a natural vertebra. Moreover, in [24], the mechanical properties are assessed at various positions of the installed vertebrae. In both cases, the comparison and visual presentation of the data are only concerned with the maximum values obtained. Comparison of the load and force processes themselves is impossible with this approach. The article [25] considers the effort and moments obtained during the robotic introduction of a needle into a human vertebra. This article also presents a graph of the increase in force as the needle is inserted. However, this graph is for a single experiment. At the same time, in studies of this type [26, 27], the possibility of noise in the received signal was minimized. In the case of a comparison of a large amount of data, such a presentation in the form of a large set of graphs [28–30] will be uninformative. Using the representation in violin diagrams would significantly increase the information content of the presented data. The limitations of this study include the use of phantoms for testing. Since the mechanical properties of medical phantoms are

close to real human tissues but are not an exact copy, the data obtained during real operations may have some differences. However, these differences are insignificant for the initial tests to verify the imaging method.

5 Conclusion

The considered method of presenting data in the form of violin diagrams makes it possible to compare data from multicomponent force-torque sensors at a qualitatively new level. The conducted experiment comparing various types of surgical interventions showed the possibility of determining the type of surgical intervention by the form of the given histogram. The data comparison experiment showed the possibility of using the violin diagram method to determine the limit values for force and moment equal to 85.4 [N] and 5.2 [Nm], respectively. In addition, an experiment with checking the possibility of displaying changes in the mechanical properties of tissues using violin diagrams made it possible to clearly see a linear increase in the values of values from 68.3 [N] for forces, 1.93 [Nm] for moments to 158.2 [N] for forces and 4.1 [Nm] for moments. The resulting visualizations provide a highly informative visual representation and allow you to compare a large number of measurements on a single graph. The histogram of the distribution of values on the diagram will allow you to determine the qualitative differences for different types of impacts, thereby highlighting the characteristic features of the signals.

Further work in this direction will consist of the implementation of the obtained method in surgical activity. The first goal of the implementation is to collect statistical data during operations. The resulting visualization method, together with the sensory system on the instruments, will allow us to evaluate the limiting forces applied by the surgeon at different stages of the operation, thereby providing information about the biomechanical properties of the patient's tissues. The obtained values will make it possible to refine the initial parameters for the design of mechatronic and robotic devices that automate surgical interventions that are currently performed manually. The second goal of implementation is the creation of systems containing surgical instruments with a sensory system and phantoms for training surgeons. Using the visualization method described above, it becomes possible to implement a convenient display for evaluating force effects and how to apply them at different stages of working with a phantom. Thus, it becomes possible to develop systems for assessing the correct operation of a surgical instrument.

Acknowledgements This study was supported by the Ministry of Health of the Russian Federation within the framework of the State Contract No. 122040100007-3 of April 1, 2022.

References

1. Abitbol J, Munir A, How J, Lau S, Salvador S, Kogan L, Kessous R, Breitner L, Frank R, Kucukyazici B, Gotlieb WH (2020) The shifting trends towards a robotically-assisted surgical interface: clinical and financial implications. Health Policy Technol, 9. https://doi.org/10.1016/j.hlpt.2020.03.003
2. Chunikhin AA, Bazikyan EA, Poduraev YV, Vorotnikov AA, Klimov DD (2019) Comparative experimental assessment of the accuracy of nanosecond laser surgery of the oral cavity when the instrument is moved by a robotic complex and a surgeon. Russian Open Med J, 8. https://doi.org/10.15275/rusomj.2019.0307
3. Liang W, Huang S, Chen S, Tan KK (2017) Force estimation and failure detection based on disturbance observer for an ear surgical device. ISA Trans, 66. https://doi.org/10.1016/j.isatra.2016.09.015
4. Fuerst D, Hollensteiner M, Schrempf A (2015) Assessment parameters for a novel simulator in minimally invasive spine surgery. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, EMBS. https://doi.org/10.1109/EMBC.2015.7319541
5. Golahmadi AK, Khan DZ, Mylonas GP, Marcus HJ (2021) Tool-tissue forces in surgery: a systematic review.https://doi.org/10.1016/j.amsu.2021.102268
6. Abed ZA, Dhhan ZTA (2022) Design and test wireless pedicle probe for spine. Pakistan J Med Health Sci, 16. https://doi.org/10.53350/pjmhs22165564
7. Molinari L, Falcinelli C, Gizzi A, Martino A (2021) Di: Biomechanical modeling of metal screw loadings on the human vertebra. Acta Mechanica Sinica/Lixue Xuebao, 37. https://doi.org/10.1007/s10409-021-01063-5
8. Solovyev M, Andrei V, Klimov D, Vladislav K, Poduraev Y (2017) Control system of the articulated arm braking mechatronic machine (AABMM). In: Annals of DAAAM and proceedings of the international DAAAM symposium. https://doi.org/10.2507/28th.daaam.proceedings.139
9. Prokhorenko L, Klimov D, Vorotnikov A, Mishchenkov D, Poduraev Y (2022) The concept of spatial motion restriction zones in a robot-assisted surgical system. J Robot Surg, 16. https://doi.org/10.1007/s11701-021-01261-0
10. Levin AA, Klimov DD, Nechunaev AA, Vorotnikov AA, Prokhorenko LS, Grigorieva EV, Astakhov DA, Poduraev YV, Panchenkov DN (2020) The comparison of the process of manual and robotic positioning of the electrode performing radiofrequency ablation under the control of a surgical navigation system. Sci Rep, 10. https://doi.org/10.1038/s41598-020-64472-9
11. Vorotnikov A, Romash E, Isaev A, Bashevskaya O, Bianchi G, Poduraev Y (2016) Uncertainty estimation of axes direction determination of industrial robot using an ellipsoid concentration model. In: Annals of DAAAM and proceedings of the international DAAAM symposium. https://doi.org/10.2507/27th.daaam.proceedings.072
12. Andrienko G, Andrienko N, Drucker S, Fekete JD, Fisher D, Idreos S, Kraska T, Li G, Ma KL, Mackinlay JD, Oulasvirta A, Schreck T, Schmann H, Stonebraker M, Auber D, Bikakis N, Chrysanthis PK, Papastefanatos G, Sharaf MA (2020) Big data visualization and analytics: Future research challenges and emerging applications. In: CEUR Workshop proceedings
13. Chen M, Hauser H, Rheingans P, Scheuermann G (2020) Foundations of data visualization. https://doi.org/10.1007/978-3-030-34444-3
14. Hanna GB, Drew T, Arnold G, Fakhry M, Cuschieri A (2008) Development of force measurement system for clinical use in minimal access surgery. Surgical endoscopy and other interventional techniques, 22. https://doi.org/10.1007/s00464-007-9489-0
15. Sorouri K, Podolsky DJ, Wang AMQ, Fisher DM, Wong KW, Looi T, Drake JM, Forrest CR (2018) Utilization of a robotic mount to determine the force required to cut palatal tissue. J Mech Behav Biomed Mater, 86. https://doi.org/10.1016/j.jmbbm.2018.06.010
16. Chanthasopeephan T, Desai JP, Lau ACW (2003) Measuring forces in liver cutting: new equipment and experimental results. Ann Biomed Eng, 31. https://doi.org/10.1114/1.1624601

17. Hou Y, Lin Y, Shi J, Chen H, Yuan W (2018) Effectiveness of the thoracic pedicle screw placement using the virtual surgical training system: a cadaver study. Oper Neurosurg, 15. https://doi.org/10.1093/ons/opy030
18. Allgaier M, Neyazi B, Preim B, Saalfeld S (2021) Distance and force visualisations for improved simulation of intracranial aneurysm clipping. Int J Comput Assist Radiol Surg, 16. https://doi.org/10.1007/s11548-021-02413-1
19. Schwalb W, Shirinzadeh B, Smith J (2017) A force-sensing surgical tool with a proximally located force/torque sensor. Int J Med Robot Comput Assis Surg, 13. https://doi.org/10.1002/rcs.1737
20. Solovyev MA, Vorotnikov AA, Grin AA, Klimov DD, Poduraev YV, Krylov VV (2020) Conceptualization of the construction of a device for measuring forces and torques during neurosurgical procedures. Biomed Eng (NY), 54. https://doi.org/10.1007/s10527-020-10002-w
21. Luengo J, García-Gil D, Ramírez-Gallego S, García S, Herrera F (2020) Big Data Preprocessing.https://doi.org/10.1007/978-3-030-39105-8
22. Meleshnikov AM, Vorotnikov AA, Klimov DD, Poduraev YV (2020) Prototype probe determining waveguide–gum contact for a robot surgical system. Russian Eng Res, 40. https://doi.org/10.3103/S1068798X20010141
23. Hollensteiner M, Furst D, Augat P, Esterer B, Schrodl F, Gabauer S, Puschel K, Schrempf A (2018) Procedure-specific validation of artificial vertebrae. IEEE Trans Biomed Eng 65:1852–1858. https://doi.org/10.1109/TBME.2017.2782797
24. Inceoğlu S, Montgomery WH, St. Clair S, McLain RF (2011) Pedicle screw insertion angle and pullout strength: Comparison of 2 proposed strategies: laboratory investigation. J Neurosurg Spine 14:670–676. https://doi.org/10.3171/2010.11.SPINE09886
25. Matsumiya K, Momoi Y, Kobayashi E, Sugano N, Yonenobu K, Inada H, Tsuji T, Sakuma I (2003) Forces and torques during robotic needle insertion to human vertebra. Int Congr Ser 1256:492–497. https://doi.org/10.1016/S0531-5131(03)00472-2
26. Vodel M, Bergelt R, Glockner M, Hardt W (2013) Synchronised data logging, processing and visualisation framework for heterogeneous sensor networks. In: Lecture notes in electrical engineering. https://doi.org/10.1007/978-3-642-28807-4_27
27. Li J, Wu X, Chen H (2011) Research on mobile digital health system based on internet of things. In: Lecture notes in electrical engineering. https://doi.org/10.1007/978-3-642-21747-0_62
28. Madaki AA, Zainon WMNW (2022) A review on tools and techniques for visualizing software requirement traceability. In: Lecture Notes in electrical engineering. https://doi.org/10.1007/978-981-16-8129-5_7
29. Meh NHBC, Sharif NABM, Harun NHB, Embong ZB (2022) Comparative analysis of image enhancement techniques based on fundus image for diabetic retinopathy. In: Lecture notes in electrical engineering. https://doi.org/10.1007/978-981-16-8129-5_113
30. Huang J, Li H, Wu H (2011) The finite element modeling for mechanical feature analysis of human lumbar L4-L5 segment. In: Lecture notes in electrical engineering. https://doi.org/10.1007/978-3-642-21762-3_49

Analysis of Hyperparameters for Workout Movements Classification Using the Convolutional Neural Network Algorithm

M. Hasyim Abdillah Pronosumarto, Jiwa Sambhuwara, S. T. Koredianto Usman, and R. Yunendah Nur Fu'Adah

Abstract After the Covid-19 pandemic hit Indonesia. Many people are starting to be aware of their own health. However, they do not understand about correct and safe workout movements. Therefore, we want to classify workout movements with a machine learning model using the Convolutional Neural Network algorithm and compare multiple hyperparameter combinations to achieve optimal outcomes, allowing developers to deploy it on websites and applications. By evaluating all hyperparameters, this research aims to optimize model performance in the domain of workout movement classification. In this study there are 22 classes containing workout movements that focus on weightlifting exercises. The dataset was obtained from a downloaded YouTube video that the author subsequently divided into multiple frames. The dataset was splitted into 90% training and 10% validation. To achieve optimal results, the authors examine the impact of batch size, learning rate, and optimizer on the accuracy parameter of the model. The dataset will be trained using InceptionV3 CNN model architecture. The result shows adaptive optimizers like Adam, RMSprop, and AdamW outperformed basic optimizers like SGD and AdaGrad, suggesting that dynamically adjusting learning rates based on gradient information is more effective. A learning rate of 0.001 paired with a high batch size yielded superior results compared to other hyperparameter combinations when training the InceptionV3 model on this dataset. We hope application developers can utilize the study's findings to improve the performance of fitness apps that involve workout movement classification.

Keywords Batch Size · Convolutional neural network · Learning rate · Optimizer · Weightlifting · Workout movements

M. H. Abdillah Pronosumarto (✉) · J. Sambhuwara · S. T. Koredianto Usman · R. Y. N. Fu'Adah
Telkom University, Bandung Regency 40257, Indonesia
e-mail: yunendah@telkomuniversity.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_21

1 Introduction

The COVID-19 outbreak has greatly affected different areas of society, including sports and physical activities. Due to public gathering limitations and the shutdown of sports facilities, people have encountered difficulties in keeping up with their fitness routines and joining organized sports events. This has prompted an increase in interest for at-home exercise and virtual sports activities. In this context, the use of Convolutional Neural Network (CNN) technology shows great potential for precise and automated labeling of sports movements, providing valuable insights for individuals and fitness professionals.

The field of machine learning, especially deep learning, has made remarkable progress in recent years and transformed several domains. Deep learning models, like CNNs, have shown impressive abilities in extracting high-level features from intricate data, which includes images and videos [1]. Large-scale datasets, computational resources, and progressive deep learning algorithms have powered these advancements. This fast-paced advancement in machine learning techniques has paved the way for more precise and efficient classification of sports movements.

However, despite progress, there is a research gap in thoroughly analyzing the relationship between specific parameters, including optimizers, learning rate, and batch size, and their impact on training CNN models for imbalanced workout movement datasets. This gap is significant because it hinders the development of robust CNN models that accurately recognize and classify various sports movements from video recordings. Most studies, like [2–5], primarily focus on analyzing individual parameters such as optimizers while keeping other hyperparameters fixed. It's important to note that the performance of a machine learning algorithm is highly dependent on its hyperparameters; however, hyperparameter optimization is not a trivial task, as it is problem-specific and can significantly impact the model's effectiveness in addressing the specific challenges of imbalanced workout movement datasets [6].

To address this gap, our research aims to study the performance of adaptive optimizers (Adam, RMSprop, and AdamW) in comparison to basic SGD and AdaGrad optimizers using various learning rates. We will also investigate how batch sizes impact accuracy and the optimal learning rate. By analyzing various parameters, our research provides valuable insights into the best setup for training CNN models on imbalanced workout movement datasets, which contributes to a more detailed understanding of this field.

The results of this study could shape future advancements in the field of sports science, empowering individuals to efficiently maintain their fitness levels. Furthermore, the research findings can inform fitness professionals when designing personalized workout programs and contribute to advancements in sports analytics. Our study aims to improve the application of CNN technology in classifying workout movements. It has potential for implementation on Android or websites, benefiting a wide range of individuals and professionals in the fitness industry.

2 Materials and Method

There are 4 steps in this research, namely collecting data, preprocessing dataset, building CNN model, and analyzing the output and the comparison. Collecting data is the longest step of this research because the collected video must be cut in the part where there is repetition of the movement. The steps taken in this study can be seen in Fig. 1.

2.1 Collecting Dataset

The dataset was gathered from two sources: YouTube videos and self-recorded videos. The videos were then converted into photos, with the number of photos per second determined according to the desired frequency. Usually, the collected videos have a duration of 6 to 8 s, ensuring sufficient coverage of each exercise movement. We utilize the scene filter feature in the VLC media player to extract video frames, resulting in a 480 × 360 photograph. While examining the video dataset, we look for diversity in various aspects such as background, actor, and camera angle to improve the generalization ability of the machine. In total, the dataset consisted of 13,853 images, with 2,734 images created by the author and 11,119 images extracted from trimmed YouTube videos. We include workout movements that are common in fitness

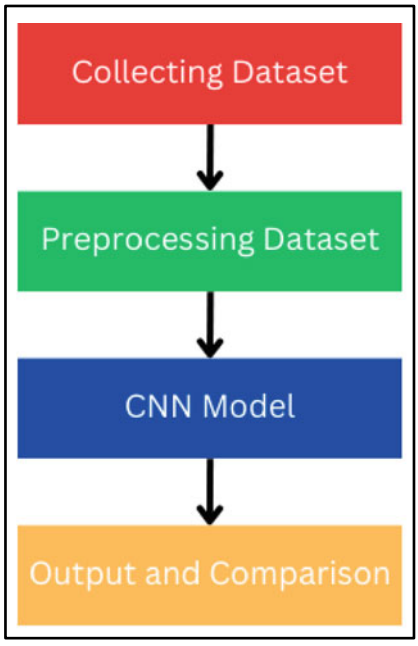


Fig. 1 Research workflow

Table 1 Dataset detail

Arms			Legs		
Barbell biceps curl	705	0	Hip thrust	557	178
Hammer curl	546	260	Leg extension	586	199
Lateral raises	843	247	Romanian deadlift	555	156
Shoulder press	512	165	Squat	742	242
Tricep dips	698	147			
Tricep pushdown	625	0			
Chest			*Back*		
Bench press	625	0	Deadlift	530	0
Chest fly machine	527	0	Lat pulldown	646	0
Decline bench press	514	334	Pull up	615	188
Incline bench press	729	0	T-bar row	668	230
Push up	601	0			
Abs					
Leg raises	514	176	**Total All**	13,853	
Plank	993	108	**Total manual**	2734	
Russian twist	522	104	**Percentage manual**	19.74%	

centers because we are targeting ordinary people who don't know the right workout movements. Table 1 shows the distribution of the dataset.

2.2 Dataset Preprocessing

In this stage, data preprocessing plays an important role in the performance of a machine learning model because the data that has been preprocessed will become cleaner and more varied as a result of the data transformations performed [7]. The main aim of preprocessing is to reduce unwanted distortion and an improvement in input data [8]. In this study the authors apply data normalization and data augmentation. Image augmentation is a technique that involves generating artificial images by applying various transformations and modifications to the original dataset, mitigating overfitting in machine learning models due to limited training data [9]. Data normalization can be done by dividing each dimension by 255 so that the relative scales of pixels become 0 to 1 [10]. The augmentations used are RandomFlip, GaussianNoise, RandomContrast, RandomBrightness, and RandomZoom. RandomFlip is used to flip photos randomly. GaussianNoise serves to add noise to the input image to match real life where the image will not always be clear. RandomContrast, Random Brightness, and RandomZoom functions to assign random amounts of contrast, brightness, and zoom to the input image to provide more complex variations [11]. The GaussianNoise used in this research is 10 and the values for RandomZoom, RandomContrast, and

RandomBrightness are 0.1. During the machine learning model training process, the dataset was partitioned into a training set comprising 90% of the data and a validation set consisting of the remaining 10%.

2.3 *Convolutional Neural Network Model*

This research uses the Inception-V3 architecture model because from several experiments with different architectures, the results of the Inception-V3 architecture are the most optimal. The loss used is categorical crossentropy because there are 22 labels used. The structure of the Inception-V3 architecture can be seen in the Fig. 2.

The basic module of Inception [24] consists of four branches concatenated in parallel: a1 × 1 kernel convolution, followed by two 3 × 3 convolutions; a 1 × 1 convolution, followed by a 3 × 3 convolution; a pooling, followed by a 1 × 1 convolution; and finally, a 1 × 1convolution. Inception consists of 10 modules, although these modules are going slightly as the net gets deeper. Five of the modules are changed with the purpose of reducing the computational cost by replacing the n × n convolutions with two convolutions, a 1 × 7followed by a 7 × 1. Two last modules replace the last two convolutions: 3 × 3 of the first branch with two convolutions each and one 1 × 3 followed by another 3 × 1, this time in parallel. In total, Inception V3 has 42 layers with parameters [13].

The author does not use the upper part of the Inception-V3 architecture so that it can adapt to the needs of the dataset. Thus, the following source code snippet exhibits an additional layer implemented by the author:

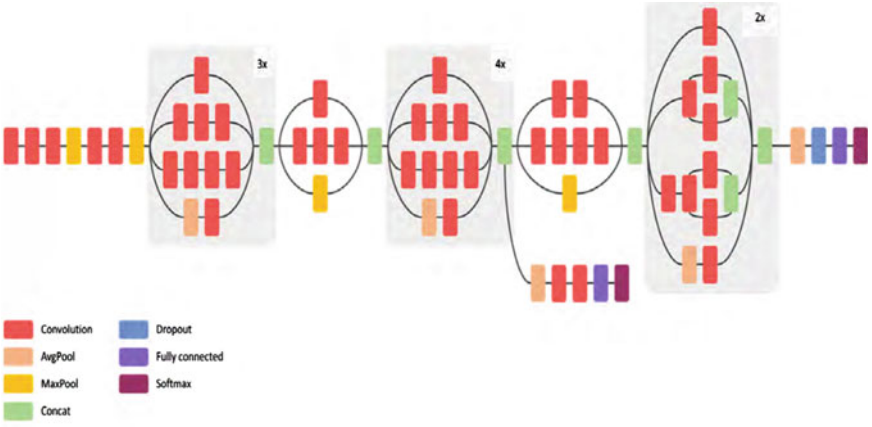


Fig. 2 Structure of Inception-V3 architecture [12]

```
x = pre_trained.output
x = tf.keras.layers.BatchNormalization(axis=-1, momen-
tum=0.99, epsilon=0.001)(x)

predictions = tf.keras.layers.Dense(len(labels), acti-
vation='softmax')(x)

workout_model = tf.keras.models.Model(inputs =
pre_trained.input,
                                       outputs = predic-
tions
                                      )
```

Adding batch normalization layer to the network and tuning the value of learning rate decay factor will improve the performance of the network by increasing the value of classification accuracy and reduce the effect of the overfitting in the learning plot [14].

CNN is a development of Multilayer Perceptron (MLP). MLP is not designed to process 2-dimensional data, therefore CNN is perfected by Yann Lechun, a researcher from AT&T Bell Laboratories in Holmdel, New Jersey, USA. Therefore, with the convolution process on CNN, the data used must have 2 dimensions, such as image and sound [15]. Figure 3 shows how CNN works.

CNN has at least 3 main layers, namely Convolutional Layer, Pooling Layer, and Fully Connected Layer. The goal of Convolutional Layer is to obtain all the high-level features of the image. The high-level features can include edges of the image too. This layer is not just limited to high-level features; it also performs an

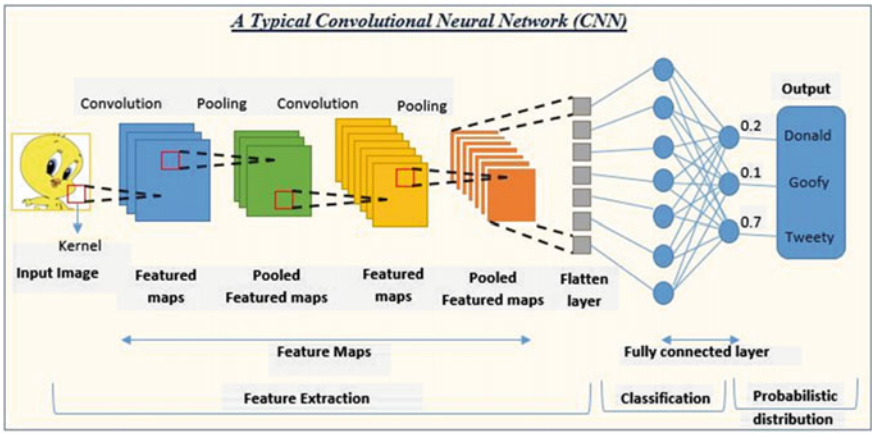


Fig. 3 How Convolutional Neural Network looks like in general [16]

operation on low-level features, such as color and gradient orientation. Pooling Layer reduces computing power for the processing of data by the method of dimensionality reduction. The main purpose of the fully connected layer is to understand and capture the complex and intricate connections between the input features and the desired output. By adjusting its weights and biases during training, this layer learns how different features relate to each other and can identify meaningful patterns in the data. This understanding allows the fully connected layer to accurately classify or predict the output based on the input data.

2.4 Analysis Method

In our paper, we utilize a rigorous data analysis methodology to derive significant insights from our image dataset. We initiate by training the dataset using a variety of hyperparameter combinations, ensuring a comprehensive exploration of the model's possible configurations. During the training process, we meticulously document the validation accuracy outcomes for every employed parameter combination at every epoch. Our objective is not only to explore different hyperparameter configurations but also to identify the most optimal combination, thereby ensuring that the training process is as efficient as possible [17]. To enable a thorough analysis, we categorize the acquired results based on our specific criteria. This categorization enables us to recognize patterns, trends, and potential correlations among the data.

Subsequently, we organize the outcomes in a systematic manner, enabling us to conduct efficient comparisons and contrasts of the effectiveness of various parameter combinations. This approach empowers us to draw meaningful conclusions and make informed decisions based on the performance of the model under different settings. To present our discoveries visually, we represent the recorded results through graphical representations. These graphs provide a clear and concise overview of the data, offering a quick and effective means of identifying key insights. By scrutinizing these plotted graphs, we can detect prevalent trends, anomalies, and latent relationships between the hyperparameters and validation accuracy, further enhancing the depth of our analysis.

3 Results

The model was trained on an imbalanced dataset with variations in hyperparameters. Specifically, five popular optimizers commonly used in CNN models—Adam, SGD, RMSProp, AdaGrad, and AdamW—were employed. Additionally, the model was trained using three different learning rates for each optimizer: 0.01, 0.001, and 0.0001. For the Adam optimizer, three different batch sizes (32, 64, and 128) were utilized, while the other optimizers maintained a consistent batch size of 64. The entire training process encompassed 20 epochs, a selection that was influenced by prior research, as

described in [18], which found this to be a suitable number of epochs for achieving convergence and model stability.

While reviewing the optimization techniques employed in this research, it is worth noting that Adadelta optimizer was not utilized. This omission is supported by prior work, as referenced in [19], which demonstrates that the Adam optimizer consistently outperforms Adadelta in terms of testing accuracy. Furthermore, our study underscores that the dynamics and optimization behavior of Stochastic Gradient Descent (SGD) are significantly influenced by the learning rate-to-batch size ratio, adding to the complexity of hyperparameter selection in training machine learning models [20, 21].

In certain research studies, it is worth noting that the Adam optimizer may not universally emerge as the optimal choice; for instance, as demonstrated in [22], customized optimizers have proven capable of delivering superior performance with considerably fewer training epochs. It's also worth noting that in this study, we only observe the training performance, and the test performance could exhibit some differences, as highlighted in [23, 24].

Our research primarily focuses on the analysis of model performance during image classification training. However, there is potential for expansion and improvement, as demonstrated in [25], where the scope extends to object detection tasks. Furthermore, it's worth noting that our study primarily focuses on the InceptionV3 model, whereas there is a wide array of CNN models, as explored in [26], that offer various avenues for further investigation.

Table 2 presents the accuracy of the model after 20 epochs of training, with a fixed batch size of 64, for different learning rates. The learning rates investigated were 0.01, 0.001, and 0.0001. The results show the performance of various optimizers commonly used in CNN models, namely Adam, SGD, RMSProp, AdaGrad, and AdamW. Analyzing the results, it is evident that the choice of optimizer and learning rate significantly impacts the accuracy of the trained model. For instance, when using a learning rate of 0.01, we observe that AdaGrad achieves the highest accuracy of 0.875762, followed closely by RMSprop with an accuracy of 0.857154. SGD and AdamW also perform reasonably well, with accuracies of 0.853465 and 0.85675329, respectively. Adam achieves a slightly lower accuracy of 0.850978. Lowering the learning rate to 0.001 resulted in improved performance for all optimizers except SGD and AdaGrad. Notably, Adam and AdamW consistently achieved higher accuracies compared to other optimizers, with values changing from 0.850978 to 0.903994 for Adam and 0.85675329 to 0.901828706 for AdamW. RMSProp also maintained a relatively high accuracy range from 0.857154 to 0.90367341. However, SGD and AdaGrad still struggled to converge, exhibiting lower accuracies of 0.853465 to 0.624398 and 0.875762 to 0.617741, respectively.

Reducing the learning rate further to 0.0001 resulted in a decrease in accuracy across all optimizers. Nonetheless, Adam, RMSProp, and AdamW still maintained higher accuracies compared to SGD and AdaGrad. Overall, these findings highlight the importance of selecting appropriate optimizers and learning rates in training models, as they significantly impact the accuracy and convergence of the trained model.

Table 2 Accuracy comparison of different optimizers with varied learning rates (Epoch 20, Batch Size 64)

Learning rate	0.01	0.001	0.0001
Adam	0.850978	0.903994	0.804219
SGD	0.853465	0.624398	0.262191
RMSprop	0.857154	0.90367341	0.808389485
AdaGrad	0.875762	0.617741	0.241177
AdamW	0.85675329	0.901828706	0.799967945

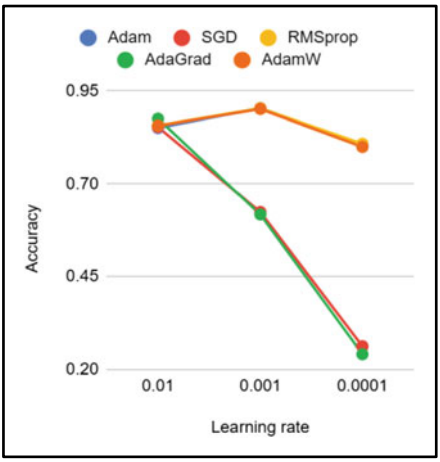


Fig. 4 Accuracy comparison of different optimizers with varied learning rates (Epoch 20, Batch Size 64)

Observing Fig. 4, it can be seen that Adam, RMSprop, and AdamW consistently perform well across all learning rates. These optimizers achieve higher accuracies compared to SGD and AdaGrad.

Adam and AdamW consistently achieve high accuracies, with slight variations based on the learning rate. RMSprop also performs well, with accuracies close to those of Adam and AdamW.

On the other hand, SGD and AdaGrad exhibit lower accuracies across all learning rates. SGD generally has better performance than AdaGrad, but it still falls behind the other optimizers. From Fig. 5, it can be observed that as the epochs progress, the accuracy generally improves for all three learning rates. For a learning rate of 0.01, the accuracy starts at 0.548364 and steadily increases, reaching 0.850978 at epoch 20. Similarly, for a learning rate of 0.001, the accuracy starts at 0.481232 and gradually improves, ultimately reaching 0.903994 at epoch 20. Comparing the two learning rates, it is evident that a learning rate of 0.001 yields consistently higher accuracies throughout the training process compared to 0.01. This indicates that a lower learning rate is more effective in this scenario. Additionally, it is worth noting that the learning rate of 0.0001 consistently maintains the lowest accuracy throughout the training, indicating that it might be too low for the model to effectively learn and

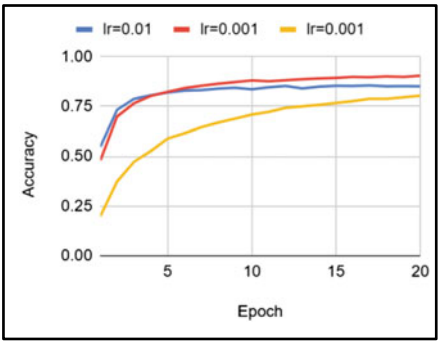


Fig. 5 Accuracy progression of model trained with Adam optimizer with varying learning rates (Batch Size 64)

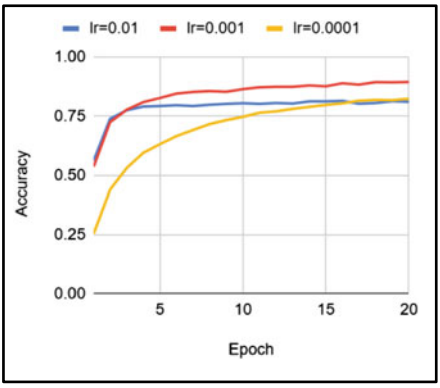


Fig. 6 Accuracy progression of model trained with Adam optimizer with varying learning rates (Batch Size 32)

improve its performance. Figure 6 presents the accuracy of a model trained using the Adam optimizer with a batch size of 32 over 20 epochs, for three different learning rates: 0.01, 0.001, and 0.0001.

During the training process, it can be observed that the accuracy steadily increases across all three learning rates as the epochs progress. At the initial epoch, the model achieves an accuracy of 0.561517 for learning rate 0.01, 0.536173 for learning rate 0.001, and 0.255935 for learning rate 0.0001. As the training continues, the accuracy improves, indicating that the model is learning and adapting to the given data. By epoch 20, the highest accuracy is achieved with learning rate 0.01, reaching 0.811758. For learning rate 0.001, the accuracy reaches 0.895653, and for learning rate 0.0001, the accuracy is 0.82427. It is evident that a learning rate of 0.001 performs the best among the three, demonstrating the model's ability to capture and generalize patterns from the data.

Figure 7 showcases the accuracy of a model trained using the Adam optimizer with a batch size of 128 over 20 epochs, utilizing three distinct learning rates: 0.01, 0.001, and 0.0001. Throughout the training process, it is evident that the accuracy steadily increases for all three learning rates as the epochs progress. Starting from the initial epoch, the model achieves accuracies of 0.529756188 for learning rate

0.01, 0.457331 for learning rate 0.001, and 0.183109 for learning rate 0.0001. As the training advances, the model continuously improves its accuracy, indicating effective learning and adaptation to the provided data. By epoch 20, the highest accuracy is attained with learning rate 0.01, reaching 0.885386586. For learning rate 0.001, the accuracy significantly increases to 0.916907, while for learning rate 0.0001, the accuracy reaches 0.766282. These results highlight the model's capability to capture and generalize patterns from the data, with lr = 0.001 delivering the most notable performance among the three learning rates.

From Fig. 8 at the beginning of training (epoch 1), the accuracy is 0.561517 for a batch size of 32, 0.548364 for a batch size of 64, and 0.529756188 for a batch size of 128. As training progresses, the accuracy generally improves for all batch sizes. On subsequent epochs, the model's performance continues to improve for all batch sizes. However, it's worth noting that the improvement rates differ. In general, the accuracy tends to increase more rapidly for larger batch sizes. For example, at epoch 5, the accuracy is 0.79315 for a batch size of 32, 0.819939 for a batch size of 64, and 0.826034665 for a batch size of 128.

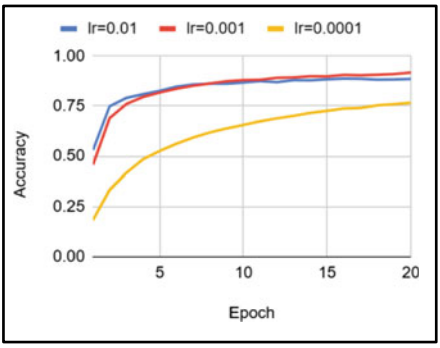


Fig. 7 Accuracy progression of model trained with Adam optimizer with varying learning rates (Batch Size 128)

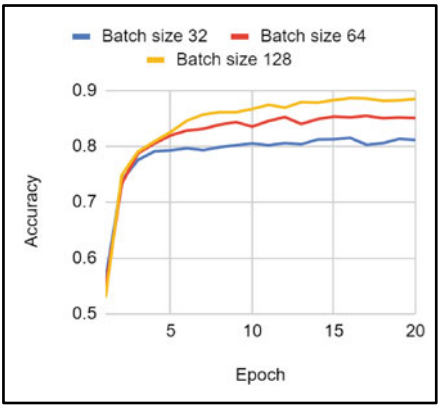


Fig. 8 Accuracy comparison across different batch sizes (32, 64, 128) using Adam optimizer with a learning rate of 0.01

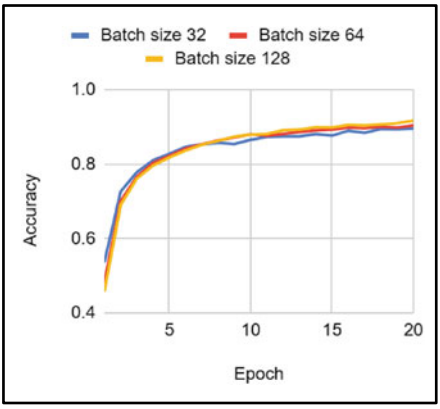


Fig. 9 Accuracy comparison across different batch sizes (32, 64, 128) using Adam optimizer with a learning rate of 0.001

This trend persists throughout training, with larger batch sizes consistently achieving higher accuracy. For instance, at epoch 20, the accuracy is 0.811758 for a batch size of 32, 0.850978 for a batch size of 64, and 0.885386586 for a batch size of 128.

On Fig. 9 in the initial epoch (epoch 1), the accuracy is 0.536173 for a batch size of 32, 0.481232 for a batch size of 64, and 0.457331 for a batch size of 128. As the training progresses, the accuracy improves for all batch sizes. As the epochs increase, the model's performance continues to improve for all batch sizes. However, similar to the previous case, the improvement rate is higher for larger batch sizes. For example, at epoch 5, the accuracy is 0.827478 for a batch size of 32, 0.822185 for a batch size of 64, and 0.817854 for a batch size of 128. This trend persists throughout training, with larger batch sizes consistently achieving higher accuracy. For instance, at epoch 20, the accuracy is 0.895653 for a batch size of 32, 0.903994 for a batch size of 64, and 0.916907 for a batch size of 128.

Figure 10 displays the accuracy achieved on each epoch using the Adam optimizer with a learning rate of 0.0001 and different batch sizes (32, 64, and 128). In the initial epoch (epoch 1), the accuracy is 0.255935 for a batch size of 32, 0.200112 for a batch size of 64, and 0.183109 for a batch size of 128. As the training progresses, the accuracy improves for all batch sizes. As the epochs increase, the accuracy continues to improve, but the rate of improvement slows down. The smaller batch sizes generally exhibit slightly higher accuracy compared to larger batch sizes. For example, at epoch 5, the accuracy is 0.631457 for a batch size of 32, 0.589188 for a batch size of 64, and 0.526067 for a batch size of 128. Throughout the training process, the accuracy steadily increases for all batch sizes, but the improvement rate is relatively slower compared to the previous cases with higher learning rates. At epoch 20, the accuracy is 0.82427 for a batch size of 32, 0.804219 for a batch size of 64, and 0.766282 for a batch size of 128.

Figure 11 illustrates the accuracy achieved on the 20th epoch using the Adam optimizer with different combinations of learning rates and batch sizes. For a learning rate of 0.01, the accuracy improves as the batch size increases. With a batch size of

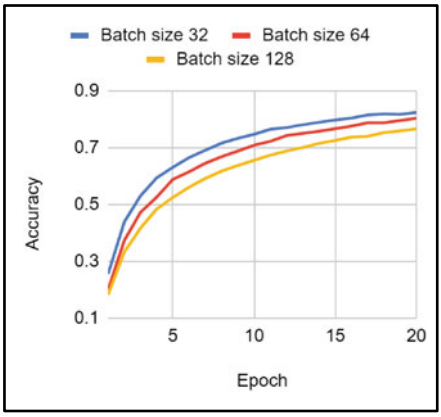


Fig. 10 Accuracy comparison across different batch sizes (32, 64, 128) using Adam optimizer with a learning rate of 0.0001

32, the accuracy is 0.811758, while it increases to 0.850978 for a batch size of 64 and further improves to 0.885386586 for a batch size of 128. This indicates that increasing the batch size allows the model to better capture patterns in the data, resulting in higher accuracy. Similarly, for a learning rate of 0.001, the accuracy also increases with larger batch sizes. The accuracy is 0.895653 for a batch size of 32, 0.903994 for a batch size of 64, and reaches the highest value of 0.916907 for a batch size of 128. This suggests that larger batch sizes contribute to improved accuracy and generalization of the model's predictions. In contrast, for a learning rate of 0.0001, the highest accuracy is obtained with a batch size of 32 (0.82427). Both smaller and larger batch sizes (64 and 128) result in lower accuracies of 0.804219 and 0.766282, respectively. This indicates that the model's performance is more sensitive to the choice of batch size when using a smaller learning rate.

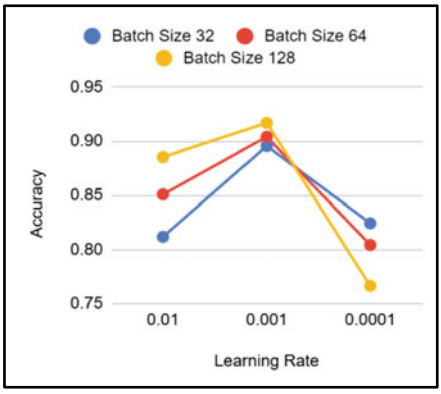


Fig. 11 Accuracy Comparison of Model Trained with Adam Optimizer at Different Learning Rates and Batch Sizes

4 Discussion

The study revealed that the adaptive optimizers (Adam, RMSprop, and AdamW) obtained higher accuracies when compared with the basic SGD and AdaGrad optimizers across all learning rates. As these optimizers dynamically adjust the learning rates based on gradient information, they are more effective for the training of Convolutional Neural Networks (CNN) models. In general, a lower learning rate of 0.001 produced better results, achieving higher accuracies when compared to the use of 0.01. Nonetheless, a learning rate of 0.0001 was deemed too small for the model to effectively learn from the given data. The model's accuracy improved across all combinations of hyperparameters during training which included various optimizers, learning rates, and batch sizes, with varying rates of improvement depending on the chosen values. Batch sizes of 64 and 128 typically resulted in faster accuracy increases compared to batch size 32. Larger batch sizes do not always ensure better performance, particularly when using lower learning rates like 0.0001. In this scenario, batch size 32 outperformed batch sizes 64 and 128, demonstrating its superior accuracy.

Compared to other related studies, our research is unique in that it analyzes multiple parameters in a comprehensive manner, using an imbalanced workout movement dataset. While certain published articles and papers have investigated the effects of specific parameters such as optimizers, batch size or learning rate, our study adopts a comprehensive perspective by examining the interplay between all of these factors. Our research investigates the performance of adaptive optimizers (Adam, RMSprop, and AdamW) compared to basic SGD and AdaGrad optimizers at different learning rates, demonstrating the superiority of adaptive optimizers in training Convolutional Neural Networks (CNN) models. Our study offers valuable insights into the optimal learning rate. It has been shown that selecting a lower learning rate of 0.001 consistently achieves higher final accuracies compared to both the higher learning rate of 0.01 and the lower learning rate of 0.0001. In addition, the analysis of batch sizes reveals that batch sizes of 64 and 128 generally lead to faster accuracy improvements. Interestingly, batch size 32 outperforms larger batch sizes (64 and 128) when using lower learning rates like 0.0001. This examination of multiple parameters contributes to a more nuanced understanding of the optimal configuration for training CNN models on imbalanced workout movement datasets.

The study offers valuable insights on the performance of adaptive optimizers, learning rates, and batch sizes when training Convolutional Neural Networks (CNN) models using an imbalanced workout movement dataset. However, there are weaknesses to consider. First, the study does not explicitly address potential biases or limitations of the imbalanced workout movement dataset used. It is important to acknowledge the representativeness and diversity of the dataset, as well as any potential sources of bias that may impact the generalizability of the findings. Second, the study does not discuss the computational resources needed to train the models with different hyperparameter configurations. Training models with larger batch sizes or more complex optimizers could have a significant computational cost that should

be considered for practical implementation. In addition, the study did not investigate the effect of other hyperparameters, such as weight decay or dropout rates, which might affect the model's performance. Moreover, the study does not give any clear insights into the interpretability of the models or the robustness of the results across datasets and various CNN architectures. Finally, the study does not mention the potential trade-offs that might exist between accuracy and other performance metrics, like training time or model complexity. This omission limits the comprehensiveness of the study's findings. Addressing these deficiencies could offer a more thorough comprehension of the boundaries and ramifications of the study's conclusions.

This study has significant implications for the training of CNN models on imbalanced workout movement datasets. The study offers valuable insights into the optimal configuration for achieving higher accuracy during training by examining the interplay between adaptive optimizers, batch sizes, and learning rates. The study's findings suggest that researchers and practitioners should carefully consider the choice of hyperparameters, as different combinations can result in varying rates of improvement. In addition, the study emphasizes the significance of selecting appropriate batch sizes because larger batches do not necessarily lead to superior performance, especially when lower learning rates are employed. These implications may provide guidance for future research and contribute to the development of more efficient training strategies for CNN models in domains with imbalanced datasets. Furthermore, this study emphasizes the necessity for additional research on the effect of hyperparameters, computational resource demands, interpretability of models, and the trade-offs between accuracy and other performance metrics. Examining these implications can enhance comprehension and application of CNN models in real-world situations.

5 Conclusion

The objective of this study was to evaluate the performance of adaptive optimizers, learning rates, and batch sizes in training Convolutional Neural Networks (CNN) models on an imbalanced workout movement dataset. The study's findings indicate that adaptive optimizers (Adam, RMSprop, and AdamW) outperform basic SGD and AdaGrad optimizers across all learning rates. Lowering the learning rate to 0.001 usually produced increased accuracies compared to 0.01, while 0.0001 was found to be inadequate for effective learning. Batches of 64 and 128 display quicker accuracy improvement than a batch of 32 at all learning rates. However, faster accuracy improvement does not inevitably correlate with higher final accuracy outcomes. Batch size 32 generates superior final accuracy outcomes at lower learning rates such as 0.0001. In the future, it is important to address the potential biases and limitations of imbalanced workout movement datasets. Additionally, exploring the computational resource requirements for different hyperparameter configurations, investigating the impact of additional hyperparameters, assessing the interpretability and robustness of models across datasets and architectures, and considering the trade-offs between accuracy and other performance metrics can lead to further advancements in training

CNN models to combat diabetes and improve healthcare outcomes. Focusing on these areas can lead to further advancements in training CNN models to combat diabetes and improve healthcare outcomes.

References

1. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444. https://doi.org/10.1038/nature14539
2. Poojary R, Pai A (2019) Comparative study of model optimization techniques in fine-tuned CNN models. In: 2019 International conference on electrical and computing technologies and applications (ICECTA), IEEE, pp 1–4. https://doi.org/10.1109/ICECTA48151.2019.8959681
3. Zohrevand A, Imani Z (2022) An empirical study of the performance of different optimizers in the deep neural networks. In: 2022 International conference on machine vision and image processing (MVIP), IEEE, pp 1–5. https://doi.org/10.1109/MVIP53647.2022.9738743
4. Chowdhury YS, Dasgupta R, Nanda S (2021) Analysis of various optimizer on CNN model in the application of pneumonia detection. In: 2021 3rd international conference on signal processing and communication (ICPSC), IEEE, pp 417–421. https://doi.org/10.1109/ICSPC51351.2021.9451768
5. Taqi AM, Awad A, Al-Azzo F, Milanova M (2018) The impact of multi-optimizers and data augmentation on tensorflow convolutional neural network performance. In: 2018 IEEE conference on multimedia information processing and retrieval (MIPR), IEEE, pp 140–145. https://doi.org/10.1109/MIPR.2018.00032
6. Fatyanosa TN, Aritsugi M (2020) Effects of the number of hyperparameters on the performance of GA-CNN. In: 2020 IEEE/ACM international conference on big data computing, applications and technologies (BDCAT), IEEE, pp 144–153. https://doi.org/10.1109/BDCAT50828.2020.00016
7. Suripto, Rahmanita N, Kirana AS, 'Teknik pre-processing Dan Classification Dalam Data Science. https://mie.binus.ac.id/2022/08/26/teknik-pre-processing-dan-classification-dalam-data-science/
8. Wankhede SS, Dondalkar S, Farkade R, Joshi A, Barapatre A, Meshram A (2022) Covid-19 detection using machine learning and deep learning. Int J Adv Res Sci Commun Technol, 764–770. https://doi.org/10.48175/IJARSCT-2951
9. Kandel I, Castelli M (2020) How deeply to fine-tune a convolutional neural network: a case study using a histopathology dataset. Appl Sci 10(10):3359. https://doi.org/10.3390/app10103359
10. Karpathy A, CS231n: convolutional neural networks for visual recognition. https://cs231n.github.io/neural-networks-2//
11. Chollet F, Omernick M, Keras documentation: working with preprocessing layers. https://keras.io/guides/preprocessing_layers/
12. Chang W-J, Chen L-B, Hsu C-H, Lin C-P, Yang T-C (2019) A Deep learning-based intelligent medicine recognition system for chronic patients. IEEE Access 7:44441–44458. https://doi.org/10.1109/ACCESS.2019.2908843
13. Pérez-Pérez BD, García Vázquez JP, Salomón-Torres R (2021) Evaluation of convolutional neural networks' hyperparameters with transfer learning to determine sorting of ripe Medjool dates. Agriculture 11(2):115. https://doi.org/10.3390/agriculture11020115
14. Ismail A, Ahmad SA, Che Soh A, Hassan K, Harith HH (2019) Improving convolutional neural network (CNN) architecture (miniVGGNet) with batch normalization and learning rate decay factor for image classification. Int J Integr Eng 11(4). https://doi.org/10.30880/ijie.2019.11.04.006
15. Eka Putra WS (2016) Klasifikasi citra Menggunakan convolutional neural network (CNN) pada Caltech 101. Jurnal Teknik ITS 5(1). https://doi.org/10.12962/j23373539.v5i1.15696

16. Shah S, Convolutional neural network: an overview. https://www.analyticsvidhya.com/blog/2022/01/convolutional-neural-network-an-overview/
17. Rochmawati N, Hidayati HB, Yamasari Y, Tjahyaningtijas HPA, Yustanti W, Prihanto A (2021) Analisa learning rate dan batch size pada Klasifikasi Covid Menggunakan deep learning dengan oOptimizer Adam. J Inf Eng Educ Technol 5(2):44–48. https://doi.org/10.26740/jieet.v5n2.p44-48
18. Hendrawan Y et al (2021) Classification of soybean tempe quality using deep learning. IOP Conf Ser Earth Environ Sci 924(1):012022. https://doi.org/10.1088/1755-1315/924/1/012022
19. Alkhalid FF (2020) The effect of optimizers in fingerprint classification model utilizing deep learning. Indonesian J Electr Eng Comput Sci 20(2):1098. https://doi.org/10.11591/ijeecs.v20.i2.pp1098-1102
20. Jastrzębski S et al (2018) Width of minima reached by stochastic gradient descent is influenced by learning rate to batch size ratio, 392–402. https://doi.org/10.1007/978-3-030-01424-7_39
21. Shi H, Yang N, Tang H, Yang X (2022) ASGD: stochastic gradient descent with adaptive batch size for every parameter. Mathematics 10(6):863. https://doi.org/10.3390/math10060863
22. Soodabeh A, Manfred V (2020) A learning rate method for full-batch gradient descent. Műszaki Tudományos Közlemények 13(1):174–177. https://doi.org/10.33894/mtk-2020.13.33
23. Narin A, Pamuk Z (2020) Effect of different batch size parameters on predicting of COVID19 cases. J Intell Syst Appl, 69–72. https://doi.org/10.54856/jiswa.202012119
24. Jepkoech J, Mugo DM, Kenduiywo BK, Too EC (2021) The effect of adaptive learning rate on the accuracy of neural networks. Int J Adv Comput Sci Appl 12(8). https://doi.org/10.14569/IJACSA.2021.0120885
25. Isa IS, Rosli MSA, Yusof UK, Maruzuki MIF, Sulaiman SN (2022) Optimizing the hyperparameter tuning of YOLOv5 for underwater detection. IEEE Access 10:52818–52831. https://doi.org/10.1109/ACCESS.2022.3174583
26. Titoriya AK, Prasad Singh M (2023) Analysis of convolutional neural network architectures for the classification of lung and colon cancer, 243–253. https://doi.org/10.1007/978-981-99-0047-3_22

Developing a Patient Module-Based Mobile Application for Effective Self-isolation Management in COVID-19 Patients



Alfi Yusrotis Zakiyyah, Eka Miranda, Meyske Kumbangsila, Mediana Aryuni, Richard, and Albert Verasius Dian Sano

Abstract The urgency of handling Covid-19 is very necessary because of its rapid spread. Self-isolation (Isoman) is a way of handling asymptomatic to mild patients. Hence, the need for mobile application technology arises to facilitate remote communication between patients and doctors. This research focuses on the development of a mobile application with a patient module that enables COVID-19 patients to effectively manage their self-isolation period. In this application, features are provided for self-reporting conditions, consultations with doctors, self-isolation monitoring and its duration, Community Health Center (Puskesmas) Hotline, tips, and tricks for undergoing self-isolation. Hence, the contribution of this study was to create a user-centric tool that empowers patients to adhere to self-isolation guidelines and addresses the communication challenges posed by physical distancing. The design of this application, utilizing the Agile Design Science Research Method (ADSRM), encompasses design and development, demonstration, evaluation, and communication. Evaluation of the application at the healthcare center is conducted using a User Experience Questionnaire (UEQ) and User Acceptance to be validated by the healthcare center. This apps used for the patient in Sawah Besar Community Health Center. The results of this development state that the health center approves of the application to help make it easier for Sawah Besar health center patients to undergo Isoman.

Keywords Apps · ADSRM · Covid 19 · Patient module · Self-isolation · User centric

A. Y. Zakiyyah (✉)
Mathematics Department, School of Computer Science, Bina Nusantara University, Jakarta 11480, Indonesia
e-mail: alfi.zakiyyah@binus.edu

E. Miranda · M. Aryuni · Richard · A. V. D. Sano
Information Systems Department, School of Systems Department, Bina Nusantara University, Jakarta 11480, Indonesia

M. Kumbangsila
Medical Doctor, Community Health Service, Sawah Besar, Jakarta, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_22

1 Introduction

Through pandemic covid 19, The Indonesian government has made continuous efforts to prevent the spread of the Covid-19 virus, such as implementing PPKM (Enforcement of Community Activity Restrictions) which was officially announced in early July 2021. However, seeing that the spread of the Covid-19 virus showed no signs of abating, the government extended the PPKM regulations and introduced new terms such as Emergency PPKM, Micro PPKM, and PPKM Levels 1–4. At its peak, Indonesia experienced its worst wave of Covid-19 cases in early July 2021, with up to 40,000 daily new cases. As a result, the PPKM measures were tightened. Subsequently, the graph of Covid-19 cases in Indonesia showed improvement, with daily cases dropping below 5,000 by the end of 2021 and currently below 1,000 in early 2022 [1, 2]. Since the emergence of Covid-19 cases in remote areas, people have become afraid to visit healthcare facilities. Individuals with mild illnesses have started to avoid seeking medical care due to the fear of contracting the virus. As a result, they resort to self-medication using online references and other sources without consulting a doctor. Telemedicine has gained popularity in developing countries in recent years, especially with the high spread of the Covid-19 virus and various social restrictions. This has shifted people's perception of telemedicine from being an alternative option to becoming their primary choice. The concept of "hospitality at home" can be applied to telemedicine, as it provides convenience for mild medical treatments to be conducted at home, reducing the risk of nosocomial infections, and saving on treatment costs [3].

Research and development's role in mobile app technology for Covid-19 prevention is highly significant. Throughout the Covid-19 pandemic, various apps with specific features have emerged to mitigate physical contact. One such example is mobile contact tracing apps used to provide information about the areas where someone has been exposed to the coronavirus, thereby enabling individuals to avoid those areas. According to Du et al.'s research [4], the technologies used in tracing apps include GPS (Global Positioning System), Bluetooth, and self-reporting users. GPS is used for location tracking and contact tracing. Shahroz et al. [5] added other technologies used in tracing apps, such as QR codes, Zigbee, Digital contract tracing with IoT, and Wi–Fi. Some of the tracing apps developed include Hamagen [6], TraceTogether [7], COVID tracker [8], Health Code [9], Smittestop [10]. The development of mobile application during pandemic also includes the vaccinatin record tracking apps, such as Indonesia's Satu Sehat [11], Malaysia's Mysejahtera [12, 13], V-Safe [14] from the United States, and TousAntiCovid from France [15].

Isabel et al.'s research examined 50 apps that emerged in Spain during the pandemic [16]. These apps were used to facilitate communication between patients and doctors. In the current situation where Covid-19 patients need to undergo recovery at home, engaging in specific activities independently without contact with others, an application to connect patients and doctors is crucial for them to interact remotely without physical meetings. Some of these apps include Mediknor [17, 18],

Coronavirus—Covid-19 [19], Ser + contra COVID19 [20], GVA Responde [21, 22], and Coronatest Navarra [23].

Like Isabel's research, Taha et al. [24] conducted a survey in Abu Dhabi regarding user satisfaction with mobile health application. The result of this study revealed a relationship between the convenience of scheduling appointments and the use of mobile health application. Another statistically significant outcome of this study was that telemedicine delivered care is thought to be on par with in person consultations in terms of quality. Additionally, it was discovered that during the COVID 19 pandemic, telemedicine was preferred over in person consultation.

In this study, we developed an app for patients at the Community Health Centre (Pusat Kesehatan Masyarakat, Puskesmas) in Sawah Besar, Indonesia. Compared to the previous study, the development of the mobile health application for self-isolation in this research certainly considers specifications such as patient requirements, patient data, and integrated healthcare facility data. This application will make it easier for patients to consult with healthcare professionals or doctors. With specific features, a patient in self-isolation (Isoman) can end their isolation period if they have fulfilled the required activities for 10 days and received approval from their doctor. The implementation of this application system will streamline and improve the management of healthcare centers in Sawah Besar, especially since the application is mobile-based, offering a more convenient and user-friendly experience compared to web-based applications.

2 Material and Methods

2.1 Related Works

Mobile health (mHealth) and electronic health (eHealth) refer to the use of technology in healthcare for remote delivery of medical services and information. The advancements in e-Health and m-Health technologies are changing the way healthcare providers and consumers interact, receive, and provide care, and access health-related information [24]. E-health is the implementation of information and communication technologies within the healthcare. The early development began in 1970 using electronic health records (EHRs) technology, and experienced rapid growth following the emergence of the movement to utilize health information technology in the economy and clinical settings, known as Health Information Technology for Economic and Clinical Health (HITECH) in 2009 [25, 26]. M-Health is the utilization of mobile devices like mobile phones, patient monitoring devices, personal digital assistants (PDAs), and other wireless devices in medical and public health practices, is rapidly becoming widespread [27]. The utilization of mobile devices in healthcare technology is highly advantageous due to their easy mobility and accessibility. The use of m-Health technology during the COVID-19 pandemic is increasing. This is to

reduce the risk of patient-doctor interactions and consequently decrease the spread of COVID-19 [28, 29].

2.2 Kodular and User Experience Questionnaire

Kodular is a website that provides tools like MIT App Inventor for creating Android applications using block programming. In other words, you don't need to manually type program code to create Android applications. Kodular offers several advantageous features, such as Kodular Store and Kodular Extension IDE (now known as AppyBuilder Code Editor), which facilitate developers in uploading Android applications to the Kodular Store and creating extension IDE block programs according to their preferences [30, 31].

In this paper, we employ the User Experience Questionnaire (UEQ) to assess the user experience with this application. Various statistical methods, including mean, variance, and standard deviation, are employed to analyze the UEQ. In statistical terms, the standard deviation serves as an indicator of the degree of variation or dispersion within a set of values. Here the standard deviation formula:

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

Suppose σ denote the population of standard deviation, N is the size of the population, x_i denote of each value from the population and μ is the population mean. Variance is the average of the squared difference from the mean, denoted by σ^2. The calculation of UEQ result begin by determining the means, variance, and standard deviation of each respondent's answer. The subsequent process involves analyzing the average and variance values grouped by categories such as attractiveness, perspicuity, dependability, efficiency, stimulation, and novelty. The final step is to compare the grouped average and variance values with the evaluation scores.

2.3 Agile Design Science Research (ADSRM)

Agile Design Science Research is a research approach that combines Agile methodology and Design Science Research (DSR) to conduct studies in information systems and related domains. It involves iterative and collaborative processes in which researchers closely collaborate with stakeholders to understand their requirements, devise solutions, and assess their effectiveness. Agile Design Science Research emphasizes flexibility, adaptability, and responsiveness to changes throughout the research journey, enabling researchers to deliver practical and applicable outcomes

that address real-world issues. By integrating design thinking, agile project management, and scientific rigor, this approach aims to produce innovative and meaningful research findings. This methodology consists of six activities: Identify Problem and Motivate, Define Objective of a Solutions, Design and Development, Demonstration, Evaluation, and Communication [12, 32]. The following are the stages of the design method using Agile Design Science Research Methodology:

A. Identify Problem and Motivate

   The initial stage of this system design is to identify the existing problem and the motivation behind the development of this system. In the scope of this research, the identified problem is the spread of the Covid-19 virus, which has limited face-to-face meetings, making it difficult to monitor Covid-19 patients who are in self-isolation under the responsibility of the Sawah Besar Community Health Center (Puskesmas). The Sawah Besar Subdistrict has an area of 6.16 km^2 with a projected population of 100,958 individuals, consisting of 5 urban villages, 49 Neighborhood Associations (RW), and 598 Neighborhood Units (RT) (Central Statistics Agency of Central Jakarta, 2021). With such area size and population density, the number of positive Covid-19 cases in the subdistrict reached more than 2,200 in January and February 2022.

B. Define Objective of a Solution

   In this research case, the objective of the determined solution is to develop a mobile-based application that bridges the gap between doctors and patients of the Sawah Besar Community Health Center to monitor the conditions of Covid-19 patients undergoing self-isolation.

C. Design and Development

   After obtaining the data and business processes at the Sawah Besar Community Health Center, we will design the information system using Unified Model Language (UML), which includes Activity Diagram, Use Case Diagram, Use Case Description, Class Diagram, and System Sequence Diagram. For the development process of the patient portal application, we will use Kodular.

D. Demonstration

   Demonstration involves the participation of users, specifically doctors and IT staff at the Sawah Besar Community Health Center, who will use this application. The purpose of conducting this demonstration is to introduce the application process, gather various opinions and feedback, and meet the specific needs of the Sawah Besar Community Health Center.

E. Evaluation

   Evaluation is conducted after the demonstration phase, gathering feedback from users and doctors at the Sawah Besar Community Health Center through techniques such as User Experience Questionnaires and User Acceptance.

F. Communication:

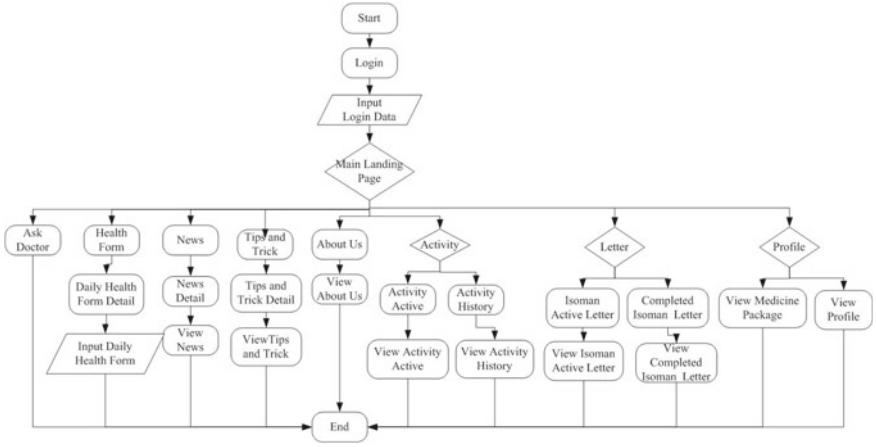


Fig. 1 Navigation structure of the self isolation application (Patient Module)

This stage aims to communicate the entire research report, including assessments and testing of the design, by publishing the results to the parties involved in the development of this application, especially the Sawah Besar Community Health Center.

Figure 1 illustrates that the design phase of this application begins with creating a design that follows the user flow from entering the application to completing the self-isolation process. The features withies this application include filling out health forms, activity history, and online consultations with doctors.

3 Result

This research utilizes several methods to analyze User Experience, including the observation method. This method is conducted by conducting online meetings with relevant parties to gather insights into the development patterns of this application. Subsequently, based on the collected data, the author will perform further analysis.

3.1 User Interface Results

The Fig. 2. below are some of the User Interface results that we have developed. The user interface (UI) of this application is being created using Kodular. The Home Page will be displayed once the patient has successfully entered the correct login code, which is provided by the doctor. Patients can choose from various features available

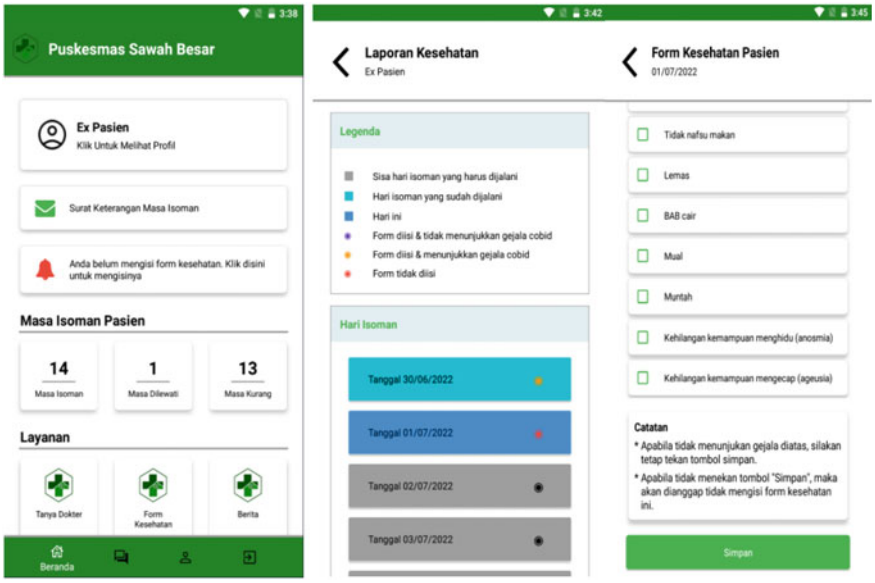


Fig. 2 User interface

in this application, such as ISoman letters, statistics, consulting with a doctor (tanya dokter), health form, news, tips and trick, about us activities, and profile.

In Fig. 3 show the self isolation letter page which contain information about the letters received from Puskesmas Sawah Besar. Figure 3a show the letter for undergoing the self isolation (Isoman) period while the letter from Fig. 3b indicating succesful or unsuccessful completion of the Isoman period.

3.2 User Acceptance

Test Analysis Results In the analysis of User Acceptance of this self-isolation patient monitoring application, the results were demonstrated to the Sawah Besar Community Health Center through a Zoom meeting held on June 22, 2022. Three participants, including doctors, medical staff, and IT personnel from the Sawah Besar Community Health Center, provided their feedback on the application. The results indicated that the application was helpful and met the needs of self-isolation patients at the Sawah Besar Community Health Center.

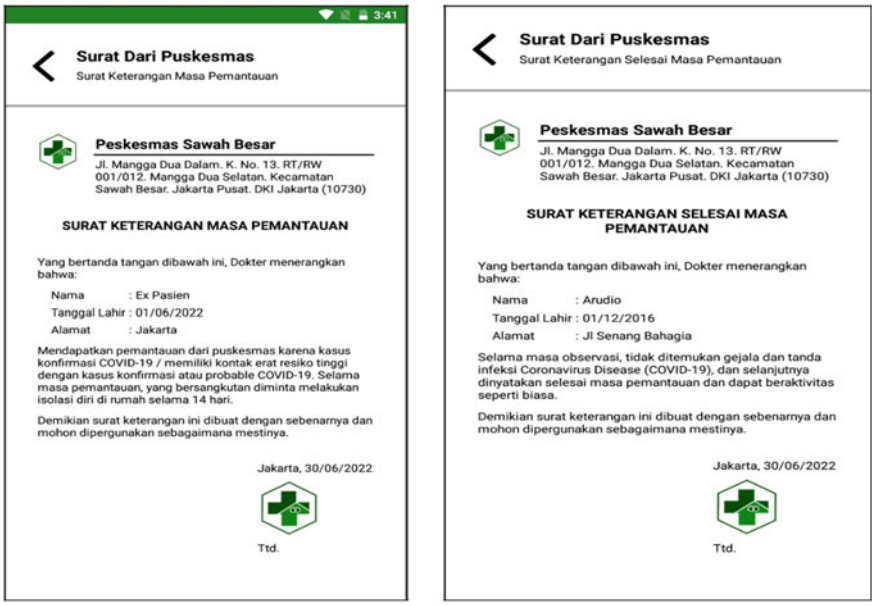


Fig. 3 Self isolation letter: **a** Self isolation active **b** self isolation is complete

3.3 User Experience Questionnaire Analysis Results (Stakeholders)

The analysis of user experience measurement using the User Experience Questionnaire (UEQ) was conducted to assess the users' experience with the application. The questionnaire was completed by our stakeholders, specifically the Sawah Besar Community Health Center, including doctors, healthcare staff, and the IT team, who had observed a demo of the application. Based on the users' evaluations, Attractiveness and Stimulation received the highest scores, indicating that the application is attractive, enjoyable, user-friendly, and comfortable to use. Additionally, the application was considered interesting, exciting, and motivating to use.

4 Discussion

We processed the UEQ data obtained from 39 respondents using the provided Data Analysis Tools by the UEQ team. We simply entered the respondents' answers into the Data Analysis Tools, which automatically calculated the results based on the provided data. In this study, we conducted a benchmark for product evaluation. Comparing the results of the evaluated product with the benchmark data enables us

Table 1 UEQ data analysis tools

Scale	Mean	Comparison to benchmark	Interpretation
Attraction	1,67	Good	10% of result better, 75% of result worse
Clearness	1,46	Above average	25% of result better, 50% of result worse
Efficiency	1,51	Good	10% of result better, 75% of result worse
Eccuracy	1,54	Good	10% of result better, 75% of result worse
Stimulus	1,69	Good	10% of result better, 75% of result worse
Novelty	1,42	Good	10% of result better, 75% of result worse

to dra conclusion about the relative quality of the evaluated product compared to others.

The Table 1 shows the results of the comparison between respondent data and UEQ Data Analysis Tools. The results indicate that 5 out of 6 scales show "Good" and 1 out of 6 scales show "Above Average". In terms of attractiveness scale, the self-isolation application we developed for Puskesmas Sawah Besar has an average score of 1,67, categorizing it as 'Good' and ranking it among the top 10% of product with the highest scores. On the clarity scale from Table 1, this self-isolation application has mean score of 1.46, categoring it as 'Above Average', indicationg that 25% of the product in the dataset scored higher, whle the remaining 50% scored lower. Moving on to the efficiency and accuracy scales, this self-isolation application has mean scores of 1.51 and 1.54, both falling into the 'Good' category, signifying that 10% of the product in the dataset scored higher than this self-isolation application, while the remaining 75% scored lower. Lastly, on the stimulation and novelty scales, this self-isolation application received a mean score of 1.69 for stimulation and 1.42 for novelty. Both scales also fall within the 'Good' category, indicating that 10% of the product in the dataset scored higher, while the remaining 75% scored lower.

5 Conclusion

Based on the research results of the patient monitoring application for self-isolation at the Sawah Besar Health Center (Patient Module) and the evaluation using Acceptance Testing and User Experience Questionnaire (UEQ) techniques, the following conclusions such as The Agile Design Science Research Methodology (ADSRM) was successfully employed in developing this patient monitoring application for self-isolation.

The evaluation results from the Acceptance Test indicate that this application will assist COVID-19 positive patients in their self-isolation period. The available features adequately cater to the patients' needs during the self-isolation process. The UEQ evaluation involving stakeholders, namely the Sawah Besar Health Center, reveals that the application is attractive, enjoyable, user-friendly, and comfortable to use. Additionally, it is considered interesting, engaging, and motivating. The results of this research can prove that the feature in the Covid—19 self-isolation patient monitoring application have practical implications, which include the potential to be send in the activities of Puskesmas Sawah Besar because the development of this application was carried out under the guidance of stakeholders (Doctors and IT personnel from the Puskesmas). TRL (Technology Readiness Level) evaluation results show that 92.5% of the application meets the criteria for being categorized as Tecnology Readiness Level (TRL) at level 5. However, the application does not meet the criteria for TRL level 6, as it has not been tested in the relevant environment, namely the Sawah Besar Health Center. For the future work, the author provides the following suggestions such as adding a live tracking feature to monitor the location of self-isolation patients. Include a feature for purchasing medications through the application. Integrate the application with the Peduli Lindungi app. These recommendations aim to enhance the functionality and usability of the application for improved self-isolation management and patient experience.

Acknowledgements This work is supported by Research and Technology Transfer Office, Bina Nusantara University as a part of BINUS Research Untuk Nusantara (BRUN) Research Grant entitled "Mobile-Based Applications for Self-Isolation of Covid-19 Patients at Puskesmas Sawah Besar" with contract number: No: 030/VRRTT/III/2023 and contract date: 01 Maret 2023. We also appreciate the support from Puskesmas Kecamatan Sawah Besar, Jakarta Pusat who help research process and our students (Rayhan and Raditya).

References

1. Krohn R, Metcalf Da, Salber P (2017) Connected health: improving care, safety, and efficiency with wearables and IoT solution. CRC Press
2. Evans RS (2016) Electronic health records: then, now, and in the future. Yearbook Med Inf 25.S 01:S48–S61
3. World Health Organization. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. Global Observatory for eHealth series, volume 3 [Internet]. 2011. [cited 2 Nov 2020]. https://www.who.int/goe/publications/goe_mhealth_web.pdf
4. Du L, Raposo VL, Wang M (2020) COVID-19 contact tracing apps: a technologic Tower of Babel and the gap for international pandemic control. JMIR Mhealth Uhealth 8(11):e23194
5. Shahroz M, Ahmad MF, Younis MS, Ahmad N, Boulos MNK, Vinuesa R, Qadir J (2021) COVID-19 digital contact tracing applications and techniques: a review post initial deployments. A review post initial deployments. Transp Eng 5
6. Golan O, Bord S, Satran C (2022) Who is following me? Public attitude towards government tracing apps in the covid Era in Israel. Europ J Public Health 32(Supplement_3)

7. Liang F (2020) COVID-19 and health code: how digital platforms tackle the pandemic in China. Soc Media Soc 6(3), [FREE Full text]. https://doi.org/10.1177/2056305120947657 [Medline: 34192023]
8. Leslie M (2020) COVID-19 fight enlists digital technology: contact tracing apps. Engineering (Beijing) 6(10):1064–1066. [FREE Full text]. https://doi.org/10.1016/j.eng.2020.09.001 [Medline: 32953197]
9. Lucivero F, Hallowell , N Johnson S, Prainsack B, Samuel G, Sharon T (2020) COVID-19 and contact tracing apps: ethical challenges for a social experiment on a global scale. J Bioeth Inq 17(4):835–839. [FREE Full text]. https://doi.org/10.1007/s11673-020-10016-9 [Medline: 32840842]
10. Sandvik KB (2020) "Smittestopp": If you want your freedom back, download now. Big Data Soc 7(2)
11. Satu sehat. https://play.google.com/store/apps/details?id=com.telkom.tracencare&hl=en&gl=US
12. Azmi PAR, Abidin AWZ, Mutalib S, Zawawi ISM, Halim SA (2022) Sentiment analysis on MySejahtera application during COVID-19 Pandemic. 2022 3rd international conference on artificial intelligence and data sciences (AiDAS), IPOH, Malaysia, pp 215–220. https://doi.org/10.1109/AiDAS56890.2022.9918748
13. Jiang X, Mohamed AE (2022) The insufficiency of the Malaysian contact tracing app from the perspective of Chinese tourists: preparing for international tourism in the post-COVID-19 world. Heliyon 8(12)
14. Liu Q, Qin C, Liu M et al (2021) Effectiveness and safety of SARS-CoV-2 vaccine in real-world studies: a systematic review and meta-analysis. Infect Dis Poverty 10:132. https://doi.org/10.1186/s40249-021-00915-3
15. Schultz É, Touzani R, Mancini J, Ward JK (2022) From contact tracing to COVID-19 pass holder; the tortured journey of the French TousAntiCovid contact tracing app. Public Health 206:5–7
16. Herrera Montano I, Pérez Pacho J, Gracia Villar S, Aparicio Obregón S, Breñosa Martinez JM, de la Torre Díez I (2022) Descriptive analysis of mobile apps for management of COVID-19 in spain and development of an innovate app in that field. Scientific Reports 12(1)
17. S.l., Mediktor TM (2022) Aplicaciones en google play. https://play.google.com/store/apps/details?id=com.teckelmedical.mediktor&hl=es&gl=US
18. Teckel Medical s.l. Mediktor. App Store. https://apps.apple.com/es/app/mediktor-eval.a-tus-s.ntomas/ id725 611856 (2022)
19. Verbaclinic. Coronavirus: Covid19. Android APKs. https://baixarapk.gratis/es/app/1506427472/coronavirus-covid 19 (2022)
20. Ser Contea AppAndAbout. Ser + contra COVID19. Aplicaciones en Google Play. https://play.google.com/store/apps/details?id=com.appanabout.defusing&hl=es&gl=US
21. Generalitad Valenciana. GVA.ES: La aplicaci.n GVA Responde supera las 45.000 descargas desde su puesta en marcha. Gabinete de
22. Comunicación Generalitad Valenciana. https://www.gva.es/es/inicio/area_de_prensa/not_detalle_area_prensa?id=881036
23. Gobierno de Navarra. CoronaTest Navarra. Aplicaciones en Google Play. https://play.google.com/store/apps/details?id=es.navarra.coronatestnavarra&hl=es&gl=US
24. Taha AR et al (2022) The integration of mHealth technologies in telemedicine during the COVID-19 era: a cross-sectional study. PLoS ONE 17(2):e0264436
25. Muhyiddin, Nugroho H (2021) A year of Covid-19: a long road to recovery and acceleration of Indonesia's development. Jurnal Perencanaan Pembangunan: The Indonesian Journal of Development Planning 5.1, 1–19.
26. Pujaningsih NY (2020) Penerapan Kebijakan Pembatasan Kegiatan Masyarakat (PKM) dalam Penanggulangan Wabah Covid-19 di Kota Denpasar. Moderat: Jurnal Ilmiah Ilmu Pemerintahan 6(3):458–470
27. Adnan MI, Pramaningtyas MD (2020) Probiotics as promising immunomodulatory agents to prevent COVID-19 infection: a narrative review. Int J Med Students 8(2)121–125

28. Chan J (2021) Exploring digital health care: eHealth, mHealth, and librarian opportunities. J Med Library Assoc JMLA 109(3):376
29. Hamsal M, Ichsan M (2021) The COVID-19 Pandemic and telemedicine adoption: challenges and opportunities. ICEBE 2020: Proceedings of the first international conference of economics, business & entrepreneurship, ICEBE 2020, 1st October 2020, Tangerang, Indonesia. European Alliance for Innovation
30. Ramya P, Padmalatha L (2022) IoT-based health monitoring system (IHMS). Machines, mechanism and robotics: proceedings of iNaCoMM 2019. Springer Singapore
31. Alda M (2020) Aplikasi CRUD Berbasis Android Dengan Kodular Dan Database Airtable. Media Sains Indonesia
32. Conboy K, Gleasure R, Cullina E (2015) Agile design science research. New horizons in design science: broadening the research agenda: 10th international conference, DESRIST 2015, Dublin, Ireland

Improvements in the Imbalanced Hemogram Data Classification



Phuoc-Hai Huynh, Ngoc-Minh Nguyen, Trung-Nguyen Tran, and Thanh-Nghi Doan

Abstract The exponential growth of hospital information systems (HIS) has led to the accumulation of vast amounts of medical data, necessitating effective analysis methods to enhance the quality and efficiency of medical services. Machine learning has emerged as a valuable technology for the automated and accurate analysis of medical data, offering potential applications in disease diagnosis and treatment. This study aims to contribute to the advancement of classification methods and address data imbalance issues in the context of hematological data. Specifically, we propose an efficient algorithm for disease classification utilizing hemogram blood test samples, employing the random forest algorithm in conjunction with the synthetic minority oversampling technique. Experimental results using real hematological data from a local hospital demonstrate the superiority of the proposed method, achieving an impressive accuracy rate of up to 97.75% and an Area Under the Curve value of up to 98.65%. The findings underscore the value of leveraging machine learning techniques in diagnoses and treatment in clinical practice, especially when integrated into HIS systems.

Keywords Machine learning · Classification · Random forest · SMOTE · Hemogram blood test · Medical informatics

P.-H. Huynh (✉) · N.-M. Nguyen · T.-N. Doan
Faculty of Information Technology, An Giang University, An Giang, Vietnam
e-mail: hphai@agu.edu.vn

Vietnam National University, Ho Chi Minh City, Vietnam

T.-N. Tran
An Giang Province Regional General Hospital, Chau Doc, Vietnam

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_23

1 Introduction

In recent years, the robust development of hospital information systems has led to a large volume of medical data that needs to be mined in order to increase the quality and efficiency of medical services. Medical data types are categories of information that are collected, stored, and analyzed in the context of healthcare. Examples of medical data types include demographic data, clinical data, laboratory data, imaging data, and genomic data. Hematological data comprises numerical or categorical values that describe the characteristics and functions of blood cells and their components. These values are obtained from hemogram blood test samples and are crucial clinical information used to diagnose, classify, monitor, and treat various diseases. Hemogram blood test samples provide valuable data for diagnosing and treating diseases, especially when many diseases have similar symptoms. Machine learning is a technology that can automatically and accurately analyze hemogram data [29]. These models are trained on a large dataset of patient blood test results and diagnoses, learning to distinguish between diseases based on the features of the blood test results. Additionally, machine learning models can detect diseases early, even when they are asymptomatic. This can improve treatment outcomes and reduce mortality rates [20]. Nowadays, more studies are being conducted to personalize treatment plans for individual patients [18]. Machine learning models can predict how patients will respond to different treatments, helping doctors choose the best treatment for each patient. By using machine learning, doctors can make faster, more accurate, and more efficient diagnoses, improving treatment outcomes for patients and reducing healthcare costs [20]. There have been numerous studies on applying machine learning models for classifying diseases in biomedical data [14, 17]. However, there has been a lack of research on hematological data [21]. In literature, many research papers are available which mainly focused on predicting diseases from hemogram data sets using data mining techniques. In [3], the authors presented a novel benchmark data set for blood diseases detection and applies several classical machine learning algorithms to classify diseases based on blood test parameters. Random Forest is proposed in [2]. This paper compares different classification algorithms using hematological data and finds that is the most effective algorithm with an accuracy of 96.47%. In addition, some studies ([1] and [11]) show that hemogram blood test samples could perform the analysis of COVID-19. Vijayarani and Sudha [24] proposed a new clustering algorithm which is named as weight based k-means algorithm is developed for identifying the leukemia, inflammatory, bacterial or viral infection, HIV infection and pernicious anaemia diseases from the hemogram blood test samples data set.

Analyzing medical data can be challenging due to imbalances in the dataset [10]. One of the main issues is that the dataset may have an unequal distribution of outcomes, with one type of outcome being more prevalent than another. This can make it difficult for machine learning algorithms to accurately predict the minority class. To address this problem, there are several approaches that can be taken [23]. The first is to collect more samples [14], but obtaining sufficient and diverse data for rare conditions remains a persistent challenge. The second approach is to use

enhanced samples generated from raw data [15]. The last approach is to use specialized techniques [16] such as oversampling, under sampling, or the use of advanced algorithms designed to handle imbalanced datasets.

The proposed research aims to address the lack of research on hematological data analysis using machine learning techniques. Although there have been studies on applying machine learning models to classify diseases in biomedical data, there is a scarcity of research specifically focusing on hematological data. The proposed research aims to address two additional gaps in the field of hematological data analysis: the limitation of accuracy in disease classification and the challenge of imbalanced data. By addressing these gaps, the research seeks to improve the efficiency and accuracy of diagnoses, enhance treatment outcomes, and reduce healthcare costs in the field of hematological diseases. In this manuscript, we present an approach for classifying diseases using hemogram blood test samples. Specifically, we have proposed random forest [6] based on the oversampling algorithm by SMOTE [8] to diagnose six diseases. This approach, which uses hemogram blood test samples as data input and combines specific features throughout the forest, is able to classify diseases with an accuracy up to 97.75%. We have collected, cleaned, preprocessed, and labelled an amount of 1766 samples from the An Giang province regional general hospital (An Giang, Vietnam) to be fed to our algorithm. However, the data collected is unbalanced that led to the classify algorithms are overfiting. So the SMOTE is suggested to address this issue. This combine approach improves the classification performance of models. Furthermore, we have used comparison analysis techniques and visualize to interpret our results and compare them with stage-of-the-art classify methods. The remainder of this paper is organized as follows. Section 2 describes material and methods. Results and discussion are explained in Sect. 3. At Sect. 4, the conclusion is given.

2 Materials and Methods

The primary goal of this research is to analyze hemogram blood test samples in order to predict diseases. The experiment was carried out in four main steps: data collection, data preprocessing, data augmentation, and classification which is explained in Fig. 1.

In the data collection step, the researchers obtained the necessary data from the laboratory information system (LIS) database of the An Giang Province Regional General Hospital. The database contained a total of 1,766 samples from patients.

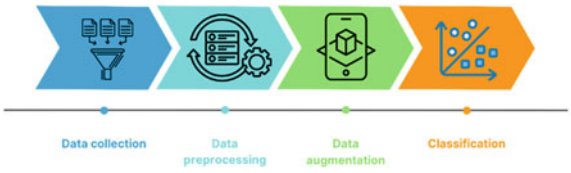


Fig. 1 The overview experiment design

After collecting the data, the next step involved data preprocessing. This step aimed to clean and prepare the data for further analysis. Outliers and missing values were removed from the dataset to ensure the quality of the data [25]. Additionally, the data was normalized to a common scale, which is often done to bring different features or variables into a similar range.

The third step involved data augmentation using the SMOTE algorithm. This technique is commonly used to address class imbalance issues in machine learning datasets. By using SMOTE, it was possible to generate new samples by interpolating between existing ones, resulting in a more accurate representation of the dataset's various classes. Finally, the preprocessed and augmented data was used for classification. Multiple machine learning algorithms were employed for this task, including support vector machines (SVMs) [9], random forests (RF) [6], neural networks (ANNs), decision trees (DT) [5], bagging of decision trees (BA) [4], gradient boosting of decision trees (GB-DTs) [13], and k-nearest neighbors (kNN) [12]. Each algorithm likely has its own strengths and weaknesses, and by comparing their performance, we can evaluate which models perform best for the given task based on the experimental results. To assess the models' performance, we used various evaluation metrics. These included classification accuracy (CA), precision (Prec), recall (Recall), F1 score, and area under the curve (AUC) [26].

2.1 Dataset and Preprocessing

In the preprocessing step of the dataset, we use many preprocessing data methods to ensure the quality and reliability of the data for further analysis. These technologies include eliminating irrelevant attributes and refilling missing values. Firstly, irrelevant attributes were eliminated from the dataset. This step involved identifying and removing any attributes that were deemed unnecessary or not relevant to the analysis. All private patient information is removed to ensure information security. The dataset was streamlined to focus on the essential features related to the complete blood count (CBC) measurements. The attributes characterize the Complete Blood Count (CBC) features as in Table 1.

We chose six common illness types for categorization in specific disease groups based on the ICD-10 codes. They include L02, D69.3, D17, J02, D50 and A94. The information of these disease code could search in website (https://icd.who.int/browse10/2010/en). In order to handle missing values, we refill missing values in the dataset using the SMOTE algorithm. Missing values can occur due to various reasons, such as data entry errors or equipment malfunctions. It is important to handle missing values appropriately to avoid biased or inaccurate results. The SMOTE technique is used to fill in the missing values based on the available data. This step ensures that all samples have complete information for the selected attributes.

Table 1 Characteristics of complete blood count (CBC) feautures

ID	Name Feature	Decription
1	Age	
2	Sex	
3	WBC (cmm)	White Blood Cell
4	RBC (million/cmm)	Red Blood Cell
5	HGB (g/dl)	Hemoglobin
6	HCT (I/I)	Hemoglobin
7	MCV (ft)	Mean Cellular Volume
8	MCH (pg)	Mean Cellular Hemoglobin
9	MCHC (g/dl)	Hemoglobin Concentration
10	PLT (/Cmm)	Platelet Count
11	LYM 1	Lymphocyte Percentage
12	LYM 2	Lymphocyte Percentage
13	MPV	Mean Platelet Volume
14	RDW	Red Cell Distribution Width

2.2 Methods

In this session, we describe the primary methods that we used to develop our efficient algorithm for disease classification based on hemogram blood test samples, including the SMOTE algorithm and Random Forest. In other methods, we use them to compare our approach.

SMOTE Algorithm. For the collected datasets, it is clear that our study faces the challenge of working with imbalanced datasets (Fig. 2). This issue is popular for medical data. Most machine learning techniques have poor performance on the minority class. One way to solve this problem is to oversample the examples in the minority class. This can be achieved by simply duplicating examples from the minority class in the training dataset prior to fitting a model. This can balance the class distribution but does not provide any additional information for the model.

SMOTE algorithms are a type of oversampling technique that are used to balance the classes in a dataset. We propose SMOTE to conjunct with classification methods to improve the accuracy of models as well as fill miss values. There are four steps in the SMOTE algorithm [8]:

Step 1: Determine the class in the dataset that has fewer instances compared to the majority class. This class is considered the minority class.

Step 2: Select a minority class instance at random to use as the foundation for creating fake samples.

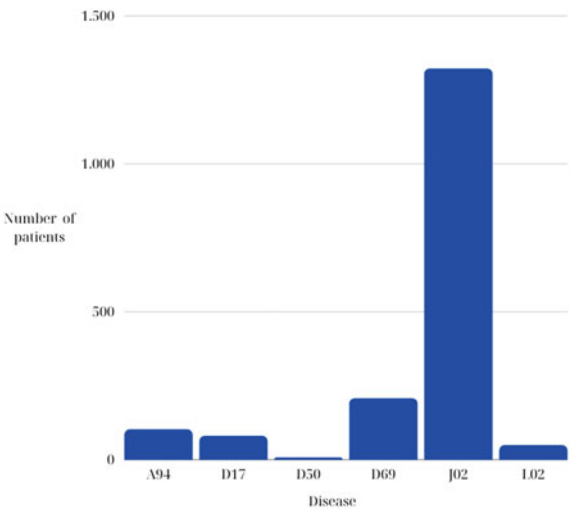


Fig. 2 Data distribution by disease label: class imbalance in the dataset

Step 3: Compute the k-nearest neighbors: Calculate the k-nearest neighbors of the selected minority class instance. The value of k is a user-defined parameter that determines the number of neighbors to consider.

Step 4: For each case of a minority class, the algorithm makes a synthetic sample by picking at random one or more of its k closest neighbors. These samples are made by interpolating the feature values between the chosen instance and its neighbors.

Random Forest Classification. Random forest is an ensemble method that combines multiple decision trees to create a more accurate and robust classifier [6]. Each decision tree is trained on a random subset of the data and a random subset of the features. The final prediction is obtained by averaging the predictions of all the trees. Random forest can handle high-dimensional and noisy data and can also provide estimates of feature importance and error rates [27]. In addition, this algorithm is effective for handling imbalanced data because the random feature selection during tree construction helps in reducing bias towards the majority class [28]. In this study, we propose a random forest to classify diseases based on hemogram blood test samples. One of the advantages of a random forest is that it can capture the nonlinear and complex relationships between the features and the target variable. Moreover, random forest is easy to implement and interpret.

3 Evaluation and Discussion

We are interested in the classification performance of our proposal (using SMOTE and Random Forest) to classify diseases based on hemogram blood test samples. As a result, we present a comparison of the classification performance between our model

and the top modern algorithms. In order to evaluate the effectiveness of classification tasks, we have implemented all algorithms in Python using the Scikit library [7].

3.1 Experiments Setup

We use a tenfold cross-validation protocol that remains the most widely used to evaluate its performance. The total classification metrics measure is used to evaluate the classification models. We train and evaluate the random forest classifier using cross-validation and various performance metrics, such as accuracy, precision, recall, F1-score, and ROC curve. In addition, we also compare our results with those of other machine learning methods, such as support vector machines (SVMs) [9], random forests (RFs) [6], neural networks (ANNs), decision trees (DTs) [19], bagging of decision trees (BA-DTs) [4], gradient boosting of decision trees (GB-DTs) [13], and k-nearest neighbors (kNN) [12]. We show that our algorithm achieves high accuracy and low error rates in classifying different diseases based on hemogram blood test samples. All experiments are run on a machine running Linux Mint, with an Intel (R) Xeon (R) CPU at 3.07 GHz, 8 cores, and 8 GB of RAM.

3.2 Classification Results

In order to demonstrate the effectiveness of the SMOTE and Random Forest, we conducted two experiments in a concise way, as follows: In the first experiment, sevven classification approaches were used for direct classification without incorporating the SMOTE algorithm. The original hemogram dataset is imbalanced and is used for training and evaluating the classification models. The purpose of this experiment was to establish a baseline performance for comparison. Table 2 and Fig. 3 show the classification results of this experiment.

Table 2 shows the performance of seven classifying models on a dataset of hemogram blood test samples. The best performing model was the RF model, which achieved a CA of 85.73%, a precision of 93.1%, a recall of 85.73%, an F1 score of 89.5%, and an AUC of 65.66%. The RF model performed better than the other models because it is able to learn complex relationships between the features and the diasease classes [22]. The other models performed well, but they did not perform as well as the RF model.

In addition, we also use AUC metric to evaluate classifying methods. The AUC metric is preferred for evaluating imbalanced data because it addresses the challenges posed by class imbalance and provides a comprehensive assessment of the model's discriminative ability, regardless of the class distribution [26]. However, the AUC metric results in Table 2 and Fig. 3 show that these classifying models face imbance issue. While the accuracy of the models is very high, the AUC is not good.

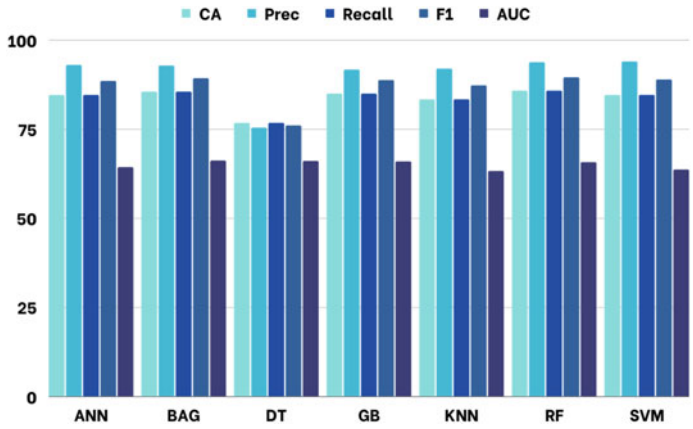


Fig. 3 Classification results (%) comparison without using SMOTE

Table 2 Comparison classification results (%) of various classifiers without using SMOTE

Model	CA	Prec	Recall	F1	AUC
ANN	84.54	92.95	84.54	88.45	64.26
BAG	85.45	92.77	85.45	89.2	66.13
DT	76.67	75.32	76.67	75.98	65.99
GB	84.88	91.63	84.88	88.71	65.85
kNN	83.3	91.88	83.3	87.2	63.19
RF	**85.73**	93.71	**85.73**	**89.5**	65.66
SVM	84.54	93.94	84.54	88.85	63.63

In the second experiment, we employed the SMOTE algorithm to address the issue of class imbalance in the dataset. By oversampling the minority class using SMOTE, we aimed to rebalance the class distribution and create a more representative training set. The augmented dataset, consisting of both the original samples and the SMOTE-generated samples, was then used for training and evaluating the classification models in this experiment. Classification results are shown in Table 3 and Fig. 4.

In Table 3, we present the comparison of various classifiers using the SMOTE algorithm for addressing class imbalances in the dataset. The classifiers evaluated in this study include SMOTE-ANN, SMOTE-BAG, SMOTE-DT, SMOTE-GB, SMOTE-KNN, SMOTE-RF, and SMOTE-SVM. Each model was trained and evaluated on the balanced dataset created using the SMOTE algorithm. The results demonstrate the performance of each classifier in terms of accuracy, precision, recall, F1-score, and AUC. Higher values indicate better performance across these metrics. The number results show that SMOTE-RF achieved a high classification accuracy of 97.75%, precision of 97.78%, recall of 97.75%, F1-score of 97.75%, and AUC of 98.65%.

Table 3 Comparison classification results (%) of various classifiers using SMOTE

Model	CA	Prec	Recall	F1	AUC
SMOTE-ANN	94.59	94.63	94.59	94.6	96.75
SMOTE-BAG	96.24	96.28	96.24	96.24	97.74
SMOTE-DT	91.22	91.3	91.22	91.24	94.73
SMOTE-GB	95.85	95.95	95.85	95.85	97.51
SMOTE-KNN	91.76	93.06	91.76	92	95.06
SMOTE-RF	**97.75**	**97.78**	**97.75**	**97.75**	**98.65**
SMOTE-SVM	94.55	94.75	94.55	94.58	96.73

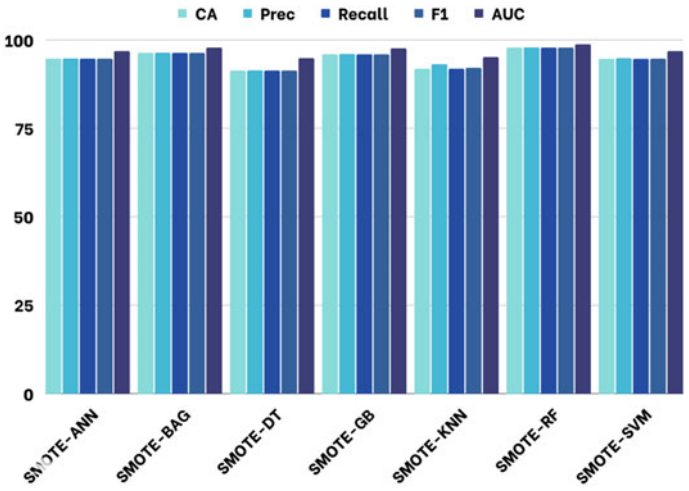


Fig. 4 Classification results (%) comparison using SMOTE

In addition, by comparing the performance metrics of different classifiers, this table provides insights into their effectiveness in handling class imbalances using the SMOTE algorithm. Table 4 can guide the selection of the most suitable classifier for the given dataset and task, considering both overall performance and specific evaluation metrics.

The provided Table 4 compares performance metrics of various machine learning models with and without using SMOTE to handle imbalanced data. The table shows that, in most cases, using SMOTE improves the performance metrics compared to the models without SMOTE. Using SMOTE generally improves classification accuracy (CA) compared to models without SMOTE, with improvements ranging from 0.46% to 6.49%. SMOTE also improves precision, indicating a better ability to correctly identify positive instances, with improvements ranging from 0.81% to 15.98%. Additionally, SMOTE enhances recall, which represents the ability of the model to correctly identify all positive instances, with improvements ranging from

Table 4 Comparison classification results (%) of various classifiers using SMOTE and without using SMOTE

Compare using SMOTE and without using SMOTE	Improve (%)				
	CA	Prec	Recall	F1	AUC
SMOTE-ANN & ANN	10.05	1.68	10.05	6.15	32.49
SMOTE-BAG & BAG	10.79	3.51	10.79	7.04	31.61
SMOTE-DT & DT	14.55	15.98	14.55	15.26	28.74
SMOTE-GB & GB	10.97	4.32	10.97	7.14	31.66
SMOTE-KNN & kNN	8.46	1.18	8.46	4.8	31.87
SMOTE-RF & RF	12.02	4.07	12.02	8.25	32.99
SMOTE-SVM & SVM	10.01	0.81	10.01	5.73	33.1

10.05% to 14.55%. The F1 score, which is the harmonic mean of precision and recall, also shows improvement with SMOTE in most cases, with improvements ranging from 4.8% to 15.26%.

While the AUC values are relatively similar between models with and without SMOTE, with slight improvements or declines observed, SMOTE tends to improve the performance of machine learning models in handling imbalanced data overall. It helps achieve better accuracy, precision, recall, and F1 score, which are crucial in accurately classifying rare or minority class instances. To compare the CA and AUC of models using SMOTE and without using SMOTE, we present the following Figs. 5 and 6.

Figure 5 shows that we observe that using SMOTE generally improves the CA of the models compared to not using SMOTE. The models consistently show higher

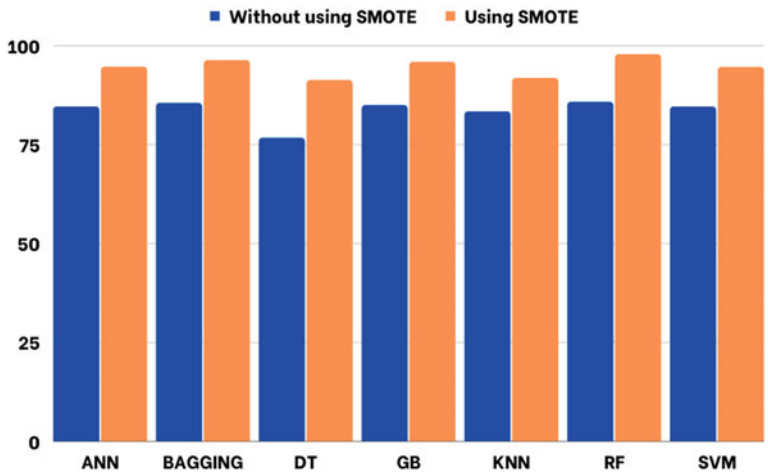


Fig. 5 Comparison CA (%) of models using SMOTE and without using SMOTE

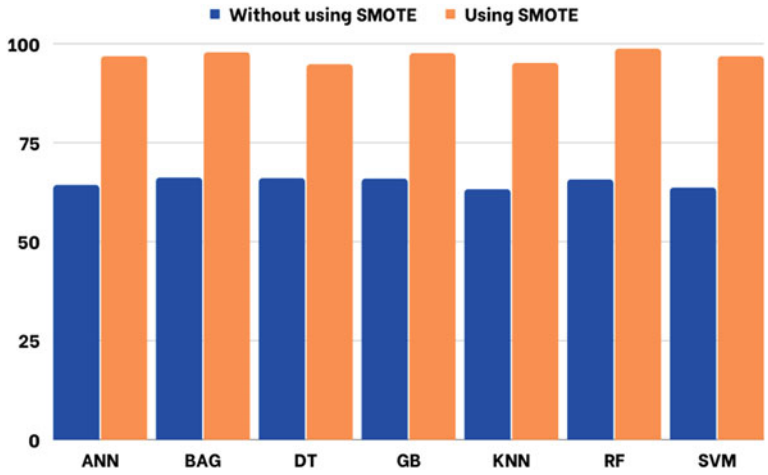


Fig. 6 Comparison AUC (%) of models using SMOTE and without using SMOTE

CA values when SMOTE is applied. For example, the RF model achieves a CA of 85.73% without SMOTE, which increases to 97.75% with SMOTE.

From Fig. 6, we can see that using SMOTE also leads to higher AUC values across the models. AUC measures the overall performance of the models in distinguishing between positive and negative instances, considering various thresholds. The AUC values show significant improvements when SMOTE is utilized. For instance, the RF model achieves an AUC of 65.66% without SMOTE, which increases to 98.65% with SMOTE. These results indicate that applying SMOTE helps to improve both the CA and AUC of the models. It effectively addresses class imbalance, allowing the models to better capture the patterns and make more accurate predictions. The substantial improvements in AUC suggest that SMOTE enhances the models' ability to rank instances correctly and distinguish between classes, leading to improved overall classification performance. Although the experimental results have very good results for the application of machine learning to hemogram blood test samples, However, one limitation of this study is that the dataset used for testing was collected from a single hospital in a specific geographic region. This may limit the generalizability of the findings to other populations or settings. Another limitation of the study is that the collected dataset only includes six diseases from the LIS database. This may limit the applicability of the findings to a wider range of diseases. Further research is needed to evaluate the proposed approach using datasets that include a more diverse range of diseases. Additionally, the study focused solely on hemogram blood test samples and did not incorporate other types of medical data. Further research is needed to evaluate the proposed approach using larger and more diverse datasets and to explore the potential of incorporating other types of medical data.

4 Conclusion

In this study, we proposed an efficient algorithm for disease classification based on hemogram blood test samples. We applied the SMOTE algorithm to address the issue of class imbalance in the dataset and used Random Forest as the classification algorithm. Our proposed approach achieved high accuracy in classifying different diseases based on hemogram blood test samples. Specifically, the approach achieved an accuracy of up to 97.75% and an AUC of up to 98.65%. The results of this study demonstrate the potential of machine learning in improving the accuracy and efficiency of disease diagnosis and treatment. By analyzing hemogram blood test samples, doctors can make faster and more accurate diagnoses, improving treatment outcomes for patients and reducing healthcare costs. Future research could focus on applying the proposed approach to larger and more diverse datasets to further validate its effectiveness. In addition, the approach could be extended to other types of medical data, such as clinical data, imaging data, and genomic data, to enable more comprehensive and personalized medical services.

Acknowledgment This research is funded by Vietnam National University Ho Chi Minh City (VNU-HCM) under grant number C2024-16-02.

References

1. Akhtar A et al (2021) COVID-19 detection from CBC using machine learning techniques. Int J Technol Innov Manag IJTIM 1(2):65–78
2. Akter F et al (2018) Classification of hematological data using data mining technique to predict diseases. J Comput Commun 6(4):76
3. Alsheref FK, Gomaa WH (2019) Blood diseases detection using classical machine learning algorithms. Int J Adv Comput Sci Appl 10:7
4. Breiman L (1996) Bagging predictors. Mach Learn 24(2):123–140
5. Breiman L et al (1984) Classification and regression T rees (Monterey, California: Wadsworth). Inc
6. Breiman L (2001) Random forests. Mach Learn 45:5–32
7. Chang C-C, Lin C-J (2011) LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol TIST 2(3):27
8. Chawla NV et al (2002) SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 16:321–357
9. Cortes C, Vapnik V (1995) Support vector machine. Mach Learn 20(3):273–297
10. Deo RC (2015) Machine learning in medicine. Circulation 132(20):1920–1930
11. Doewes RI et al (2022) Diagnosis of COVID-19 through blood sample using ensemble genetic algorithms and machine learning classifier. World J Eng 19(2):175–182
12. Fix E, Hodges J (1952) Discriminatory analysis-nonparametric discrimination: Small sample performance. California Univ, Berkeley
13. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat, 1189–1232
14. Huynh P-H et al (2021) Enhancing COVID-19 prediction using transfer learning from Chest X-ray images. In: 2021 8th NAFOSTED conference on information and computer science (NICS), pp. 398–403. IEEE

15. Huynh P-H et al (2019) Enhancing gene expression classification of support vector machines with generative adversarial networks. J Inf Commun Converg Eng 17(1):14–20
16. Huynh P-H et al (2020) Improvements in the large p, small n classification issue. SN Comput Sci 1:1–19
17. Huynh PH, Nguyen VH (2023) A novel ensemble of support vector machines for improving medical data classification. Eng Innov 4:47–66
18. Kalantari A et al (2018) Computational intelligence approaches for classification of medical data: state-of-the-art, future challenges and research directions. Neurocomputing 276:2–22
19. L Breiman RAO, J Friedman CJ (1984) Stone: classification and regression trees. Wadsworth Int Group 8:452–456
20. MacEachern SJ, Forkert ND (2021) Machine learning for precision medicine. Genome 64(4):416–425
21. Obstfeld AE (2023) Hematology and machine learning. J Appl Lab Med 8(1):129–144
22. Qi Y (2012) Random forest for bioinformatics. Ensemble Mach Learn Methods Appl, 307–323
23. Rahman MM, Davis DN (2013) Addressing the class imbalance problem in medical datasets. Int J Mach Learn Comput 3(2):224
24. Vijayarani S, Sudha S (2015) An efficient clustering algorithm for predicting diseases from hemogram blood test samples. Indian J Sci Technol 8(17):1
25. Vinisha FA, Sujihelen L (2022) Study on missing values and outlier detection in concurrence with data quality enhancement for efficient data processing. In: 2022 4th international conference on smart systems and inventive technology (ICSSIT), pp 1600–1607 IEEE
26. Vujović Z (2021) Classification model evaluation metrics. Int J Adv Comput Sci Appl 12(6):599–606
27. Wang Q et al (2018) An efficient random forests algorithm for high dimensional data classification. Adv Data Anal Classif, 1–20
28. Zhu M et al (2018) Class weights random forest algorithm for processing class imbalanced medical data. IEEE Access. 6:4641–4652
29. Zini G (2005) Artificial intelligence in hematology. Hematology 10(5):393–400

Diagnose Skin Face Problems by Comparing Classification Algorithms



Marsya Ardini, Alzha Rizqie Kinanta, Vincensius Bunni Palagoro, Michael Alessandro Kevin Wibowo, and Aripin

Abstract Data mining is a process of analyzing a set of existing data in the database so that information is obtained that is used in the next stage. Classification is a technique by forming models from unclassified data, to be used to classify new data. Classification is included in the type of supervised learning, meaning that training data is needed to build a classification model. Each classification has a choice of algorithms, which are often used are the naive Bayes algorithm, k-nearest neighbor, decision tree and support vector machine. In this study, a comparison was made to a case study of decision making in skin cases. The research method used in this study is a training method using data training and testing mode using data testing with a comparison of training and testing 70%: 30%, 80%: 20%, 90%: 10% by naive bayes algorithm, K-nearest neighbor, decision tree. Evaluate the training and testing carried out on the naive Bayes classification algorithm, nearest neighbor, decision tree in this study, namely how much accuracy the algorithm produces. The results of the comparison of the nearest neighbor, naive Bayes and decision tree classification algorithms used in the case study of clothing pattern selection decision making state that the decision tree classification algorithm is the classification algorithm that has the highest level of accuracy, reaching 97.5% in testing.

Keywords Naive Bayes · Nearest neighbor · Decision tree · And support vector machines

1 Introduction

Data mining involves examining a database's data set to extract information for subsequent stages, employing a method to derive patterns from the data. Mirkin defines data mining as the process of seeking patterns within a dataset stored in a

M. Ardini (✉) · A. R. Kinanta · V. B. Palagoro · M. A. K. Wibowo · Aripin
Biomedical Engineering, Faculty of Engineering, Universitas Dian Nuswantoro, Semarang, Indonesia
e-mail: 513202100158@mhs.dinus.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_24

database, intending to analyze it later, and generate specific information for subsequent processes [1]. One methodology for data analysis is classification [1, 2], which constructs a model from unclassified data samples to categorize new samples into similar classes. Classification, a form of supervised learning, entails analyzing a dataset first and then using the identified patterns to classify test data. The process encompasses learning and classification stages. During learning, a classification algorithm is applied to analyze training data, and in classification, testing data is used to verify the accuracy of the employed classification [2, 3]. Classification techniques fall into four categories based on distinct mathematical concepts: statistics, distance-based, decision tree-based, and network-based. In this research, the researcher aims to compare the naive Bayes classification algorithm, nearest neighbor classification, and decision tree using MATLAB [4], focusing on their respective accuracy levels. The chosen case study revolves around the decision-making process for pattern selection on the skin, where the data is presented as cases with class labels such as normal skin, acne-prone skin, and skin with large pores [5]. A case refers to a compilation of information in the form of criteria used to make decisions regarding pattern selection on the skin [6].

2 Research Methods

2.1 Data Mining

Data mining, as defined by various sources [1, 7], is the extraction of valuable knowledge or information from extensive datasets. Its objective is to uncover patterns, relationships, and insights within data, facilitating more informed decision-making. The process involves the application of statistical analysis methods, artificial intelligence, and mathematical modeling techniques to scrutinize and extract information from the data [8]. While data mining and knowledge discovery in databases (KDD) are often used interchangeably to denote the extraction of hidden information from large databases [9], they encompass different concepts with a mutual relationship. Within the overall KDD process, data mining represents a distinct stage. The KDD process unfolds as follows:

Remark 1 Before delving into the information extraction phase of Knowledge Discovery in Databases (KDD), it is imperative to undertake the selection of data from a pool of operational data. This chosen subset of data, earmarked for the data mining process, is subsequently stored in a distinct file, separate from the operational database [10].

Remark 2 Prior to initiating the data mining process, a crucial preliminary step involves cleansing the data that serves as the focal point of Knowledge Discovery in Databases (KDD) [11]. This cleansing process encompasses tasks like eliminating

duplicate data, verifying data consistency, and rectifying errors such as typographical mistakes.

Remark 3 Coding serves as a transformation applied to the selected data, rendering it well-suited for the subsequent data mining process. This coding phase within Knowledge Discovery in Databases (KDD) is inherently creative, relying heavily on the specific type or pattern of information sought within the database [12].

2.2 Matlab

Matlab, short for Matrix Laboratory, stands out as a robust, matrix-oriented computing environment and programming language crafted by MathWorks. With capabilities spanning mathematical calculations, data analysis, visualization, and simulation modeling, Matlab operates within its own programming language specifically tailored for mathematical calculations and data analysis purposes [13]. This language employs a syntax akin to traditional mathematics, enriched with a plethora of mathematical functions and a toolbox designed to execute diverse tasks—ranging from signal processing, image processing, statistics, optimization to simulation modeling, and beyond [14, 15]. This toolbox conveniently offers pre-built functions and algorithms for various applications.

2.3 Classification

Classification is a data mining technique that generates models and is employed to predict the category or class of objects within a database. The classification process involves two key phases: the classification stage and the learning stage [16]. During the learning stage, classification algorithms construct a classification model through the analysis of training data [17]. The learning stage can be viewed as the process of establishing a function or mapping $Y = F(X)$, where Y represents the predicted class, and X denotes the tuple for which you wish to predict the class. It entails learning a target function that maps each set of attributes x to one of the predefined class labels Y. This learning process utilizes training data, and in the case of the Naive Bayesian Classifier, it involves calculating probability values during the learning phase. Some of the data with known class labels are used to create an approximate model. The model that has been created is then tested with other data to assess its accuracy [18].

Naive Bayes. The Naive Bayesian Classifier is a straightforward classification approach that relies on the utilization of Bayes' Theorem, assuming that the explanatory variables are independent of each other [19]. The assumption of attribute independence is seldom observed in real data. However, even when this assumption is violated, the Naïve Bayes classification still exhibits a relatively strong performance,

as demonstrated in multiple empirical studies. The presence of one attribute (variable) is unrelated to the presence of other attributes (variables). This algorithm forecasts future probabilities by drawing from past experiences.

The basis of the naive theorem used in programming is the Bayes formula:

$$P(A|\mathrm{B}) = (P(B|\mathrm{A}) * \frac{\mathrm{P(A)}}{\mathrm{P(B)}}) \tag{1}$$

This means that the probability that A will occur as B is determined by the probability that B occurs when A, the probability A, and the probability B.

K-Nearest Neighbor. KNN is a classification method derived from data that utilizes data-driven learning. It is particularly employed in supervised learning, where instances are classified by identifying the most common category among their nearest neighbors within the KNN framework [19, 20]. Determining the smallest distance between the data to be assessed and the K closest neighbors in the training data [21]. The most number of classes with the shortest distance will be the class where the evaluation data is located. The steps for the classification of the KNN algorithm are as follows:

1. Find the value for the parameter K, which corresponds to the count of the closest neighbors.
2. Calculate the distance between the training data and new data.
3. Determine the nearest neighbors based on the kth minimum distance and yhen sort.
4. Check the class from the nearest neighbor.
5. Use a simple majority from the nearest neighbor class as the new data prediction value [22].

The basis of the K-nearest neighbors theorem used in programming is the Bayes formula:

$$dis(x_1, x_2) = \sqrt{\sum_{i=0}^{n} (x1i - x2i)^2} \tag{2}$$

Decision Tree. A decision tree is a predictive model that employs a tree-like or hierarchical structure. The fundamental idea behind a decision tree is to convert data into a tree representation and rules for making decisions. This tree is a data structure that takes on the appearance of a tree, comprising a sequence of nodes and connections [4]. In a tree, there are three different types of nodes, namely root nodes, branch/internal nodes, and leaf nodes, all connected through vectors. Each node has the ability to have zero or more child nodes. The decision tree serves as a simplified illustration of a classification method for a finite set of classes. In this representation, attribute names are assigned to internal and root nodes, the edges are annotated with potential attribute values, and leaf nodes are distinguished by various class labels

[23]. The highest node in a decision tree is called the root, and usually represents the attribute that has the greatest impact on a particular class. In general, Decision Tree follows a top-down search approach to reach a solution. In the classification process, the attributes will be tested by following the path from the root node to the leaf node, and thus, a new class will be classified.

Support Vector Machines. Support Vector Machine (SVM) is a machine learning technique applied for making predictions in both classification and regression scenarios [24]. SVM follows the fundamental principle of a linear classifier, which is primarily suitable for situations where classification can be achieved through linear separation. However, SVM has evolved to handle non-linear challenges by introducing the concept of kernels within high-dimensional spaces. It excels in solving classification problems, especially when dealing with datasets containing a vast number of features. This approach involves constructing a model that distinguishes two distinct classes by optimizing the margin, defined as the distance between the nearest data points from the two classes. SVM's operation revolves around the search for a hyperplane, a dividing plane that maximizes the separation between the two classes, pushing the closest data points as far apart as possible.

3 Results

3.1 Naive Bayes

Algorithm Training. The method used in this study is a training method using training data with a comparison of training 70%:30%, 80%:20%, 90%:10%. Evaluate the training carried out on the naive Bayes classification algorithm.

The comparison of the resulting accuracy levels can be seen in Table 1. And it can be seen from the table, the highest accuracy value is obtained at a percentage split of 70%, namely 76, 19% which means that this experiment has high accuracy. Likewise for other percentage splits, an accuracy value of above 50% is obtained, this means that this method has high accuracy.

Algorithm Testing. The method used in this study is a testing method using testing data with a comparison of testing 10%:90%, 20%:80%, 30%:70%. Evaluate the testing carried out on the naive Bayes classification algorithm.

Table 1 Results of training naive bayes data accuracy

Training data	Accuracy
90	66.67
80	66.67
70	76.19

Table 2 Results of naive bayes testing data accuracy

Training data	Accuracy
10	83
20	75
30	67

The results of the comparison of the resulting accuracy levels can be seen in Table 2. As it can be seen from the table, the highest accuracy value was obtained at a 10% percentage split, namely 83%, which means that this experiment has high accuracy. Likewise for other percentage splits, an accuracy value of above 50% is obtained, this means that this method has high accuracy.

3.2 Nearest Neighbor

Training The Nearest. The method used in this study is a training method using training data with a comparison of training 70%:30%, 80%:20%, 90%:10%. Evaluate the training carried out on the K-Nearest Neighbor classification algorithm.

The resulting comparison of the accuracy levels can be seen in Table 3. And it can be seen from the table, the highest accuracy value was obtained at a 70% percentage split, namely 80.95%, which means that this experiment has high accuracy.

Algorithm Testing. The method used in this study is a testing method using testing data with a comparison of testing 10%:90%, 20%:80%, 30%:70%. Evaluate the testing carried out on the K-Nearest Neighboar classification algorithm.

The results of the comparison of the resulting accuracy levels can be seen in Table 4. And it can be seen from the table, the highest accuracy value is obtained in the percentage split 10% that is 100% which means that this experiment has high accuracy. The percentage of testing at 20% and 30% obtained a low accuracy value, which means that the test using the split percentage has a low level of accuracy.

Table 3 Data accuracy results of training nearest neighbor

Training data	Accuracy
90	77.78
80	75
70	80.95

Table 4 Data accuracy results of testing nearest neighboar

Training data	Accuracy
10	100
20	41.67
30	55.5

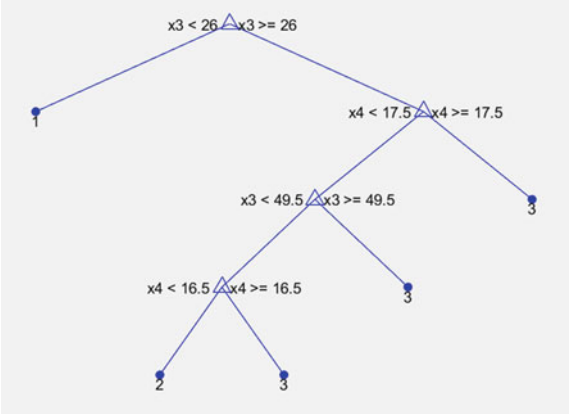


Fig. 1 Results of the accuracy of the decision tree classification

3.3 Decision Tree

The nodes in a tree are divided into three, namely root nodes, branch/internal nodes and leaf nodes which are interconnected. The nodes are connected by a vector. The decision tree will look for solutions to problems by making criteria as nodes that are interconnected to form a tree-like structure. The results of this test mode will achieve a large accuracy rate of up to 97.5% (Fig. 1).

4 Discussion

4.1 Naive Bayes

Algorithm Training. In the classification algorithm naive Bayes is a statistical calculation by calculating the probability of similarity of old cases on a case basis with new cases. This model will be tested with the percentage split training mode with the percentage of training data as much as 90, 80, 70. The results of the comparison of the resulting accuracy levels can be seen in table I. And it can be seen from the table, the highest accuracy value is obtained at a percentage split of 70%, namely 76, 19% which means that this experiment has high accuracy. Likewise for other percentage splits, an accuracy value of above 50% is obtained, this means that this method has high accuracy.

Algorithm Testing. This model will be tested using the percentage split test mode with the percentage of test data as many as 10, 20, 30. The results of the comparison of the resulting accuracy levels can be seen in Table II. And it can be seen from the table, the highest accuracy value was obtained at a 10% percentage split, namely 83%, which means that this experiment has high accuracy. Likewise for other percentage

splits, an accuracy value of above 50% is obtained, this means that this method has high accuracy.

Output Class of Test Results

1. Split dataset for the training and testing process with the proportion of 90%:10%. Testing with acne classes produces pictures of acne-prone skin. In this test the system can read the results correctly, this proves that the experiment has the right level of accuracy.
2. Split the dataset for the training and testing process with the proportion of 80%:20%. Testing with large pore class results in skin images with large pores. In this test the system can read the results correctly, this proves that the experiment has the right level of accuracy.
3. Split the dataset for the training and testing process with the proportion of 70%:30%. Testing with a normal skin class produces an image of normal skin. In this test the system can read the results correctly, this prove that the experiment has the right degree of accuracy.

4.2 Nearest Neighbor

Training the Nearest

Neighbor algorithm is a classification based on analogy, namely comparing test data with training data that is close to and has a resemblance to the test data. This model will be tested using the percentage split training mode with training data percentages of 90, 80, 70. The resulting comparison of the accuracy levels can be seen in Table III. And it can be seen from the table, the highest accuracy value was obtained at a 70% percentage split, namely 80.95%, which means that this experiment has high accuracy.

Algorithm Testing

This model will be tested using the percentage split testing mode with the percentage of test data as many as 10, 20, 30. The results of the comparison of the resulting accuracy levels can be seen in Table 4. And it can be seen from the table, the highest accuracy value is obtained in the percentage split 10% that is 100% which means that this experiment has high accuracy. The percentage of testing at 20% and 30% obtained a low accuracy value, which means that the test using the split percentage has a low level of accuracy.

4.3 Decision Tree

The decision tree algorithm is an algorithm that is commonly used for decision making. Tree is a data structure that in shape resembles a tree, which consists of a

series of nodes (vertices) and edges (edges). The nodes in a tree are divided into three, namely root nodes, branch/internal nodes and leaf nodes which are interconnected. The nodes are connected by a vector. The decision tree will look for solutions to problems by making criteria as nodes that are interconnected to form a tree-like structure. The results of this test mode will achieve a large accuracy rate of up to 97.5%.

5 Conclusion

1. The results of the comparison of the nearest neighbor, naive Bayes and decision tree classification algorithms used in the case study of clothing pattern selection decision making state that the decision tree classification algorithm is the classification algorithm that has the highest level of accuracy, reaching 97.5% in testing.
2. As for using the percentage split testing mode which has the highest level of accuracy, namely the nearest neighborboar classification algorithm with a proportion of 10%, which reaches 100%. Meanwhile, at the proportion of 20% and 30%, inaccurate results were obtained, namely below 50% because the hue,saturation, value, and area methods were not quite right in this case.
3. In the Naive Bayes method experiment, the portion of training and testing data is 90:10%. 80:20%. 70:30% has an accuracy value above 50% which means that in this experiment the level of accuracy is high.

References

1. Verma AK, Pal S, Kumar S (2019) Comparison of skin disease prediction by feature selection using ensemble data mining techniques. Inform Med Unlocked 16. https://doi.org/10.1016/j.imu.2019.100202
2. Ramageri BM, Data mining techniques and applications
3. Comparison of several classification algorithms on skin problem training and testing case study face
4. Sartika D, Sensuse DI (2017) Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian
5. Choy SP et al (2023) Systematic review of deep learning image analyses for the diagnosis and monitoring of skin disease. npj Digital Med 6(1). https://doi.org/10.1038/s41746-023-00914-8
6. Teramoto A et al (2019) Automated classification of benign and malignant cells from lung cytological images using deep convolutional neural network. Inform Med Unlocked 16. https://doi.org/10.1016/j.imu.2019.100205
7. Mirkin BG, Mirkin B, Data analysis, mathematical statistics, machine learning, data mining: similarities and differences. https://www.researchgate.net/publication/254048160
8. Jurnal H, Ali Ridla M, Artikel Abstrak I (2022) Perbandingan Algoritma Pembelajaran Mesin untuk Klasifikasi Warna Kulit Berdasarkan Warna Piksel Citra 1(1):33–41

9. Eska (2016) Data mining for predicting wallpaper sales using the C45 algorithm. JURTEKSI (Jurnal Teknol dan Sist Inf) 2:9–13
10. Balaji VR, Suganthi ST, Rajadevi R, Krishna Kumar V, Saravana Balaji B, Pandiyan S (2020) Skin disease detection and segmentation using dynamic graph cut algorithm and classification through Naive Bayes classifier. Meas J Int Meas Confed 163:107922. https://doi.org/10.1016/j.measurement.2020.107922
11. Goswami T, Dabhi VK, s HB (2020) Skin disease classification from image—a survey. 2020 6th international conference advance computer communication system ICACCS 2020, pp 599–605.https://doi.org/10.1109/ICACCS48705.2020.9074232
12. Wan L et al (2022) Detection algorithm for pigmented skin disease based on classifier-level and feature-level fusion. Front Public Health 10. https://doi.org/10.3389/fpubh.2022.1034772
13. Witten IH, Frank E, Hall MA, Pal CJ (2016) Data mining: practical machine learning tools and techniques (4th ed.). Morgan Kaufmann
14. Wan L, Ai Z, Chen J, Jiang Q, Chen H, Li Q, Lu Y, Chen L (2022) Detection algorithm for pigmented skin disease based on classifier-level and feature-level fusion. Front Public Health 20(10):1034772. https://doi.org/10.3389/fpubh.2022.1034772.Erratum.In:FrontPublicHealth.2023Jul18;11:1229178.PMID:36339204;PMCID:PMC9632750
15. Qin Z, Liu Z, Zhu P, Xue Y (2020) A GAN-based image synthesis method for skin lesion classification. Comput Methods Programs Biomed 195:105568. https://doi.org/10.1016/j.cmpb.2020.105568
16. Mohammed SS, Al-Tuwaijari JM (2021) Skin disease classification system based on machine learning technique: a survey. IOP Conf Ser Mater Sci Eng 1076(1):012045. https://doi.org/10.1088/1757-899x/1076/1/012045
17. Debelee TG (2023) kin lesion classification and detection using machine learning techniques: a systematic review. Diagnostics 13(19). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/diagnostics13193147
18. Pujol FA, Pujol M, Jimeno-Morenilla A, Pujol MJ (2017) Face detection based on skin color segmentation using fuzzy entropy. Entropy 19(1). https://doi.org/10.3390/e19010026
19. Harto B (2019) Comparison of data mining methods using the naïve Bayes algorithm and K-nearest neighbor in predicting immunotherapy success. https://jurnal.buddhidharma.ac.id/index.php/te
20. Hatem MQ (2022) Skin lesion classification system using a K-nearest neighbor algorithm. Vis Comput Ind Biomed Art 5(1). https://doi.org/10.1186/s42492-022-00103-6
21. Boiculese VL, Dimitriu G, Moscalu M (2013) Improving recall of k-nearest neighbor algorithm for classes of uneven size. In: 2013 E-Health and bioengineering conference, EHB 2013. https://doi.org/10.1109/EHB.2013.6707403
22. Sharma U, Vidyapeeth B (2020) Fake news detection using machine learning algorithms. https://www.researchgate.net/publication/343916946
23. Rosela Y (2019) Implementation of decision tree classification to analyze the status of sales of goods using C4.5 (Case Study: Pt Mata. hari Department Store Medan Mall). J Pelita Inform 18(1):143–150
24. Kaur P, Singh G, Kaur P (2019) Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification. Inform Med Unlocked 16. https://doi.org/10.1016/j.imu.2019.01.001

Quickness Aspect Talent Identification (QATI) System Based on the Internet of Things



Rohmat Tulloh, Jordy Marchelino Lumban Gaol, Dery Rimasa, and Asep Mulyana

Abstract Identifying an athlete's speed and quickness talent is a crucial stage in the world of sports that can determine an athlete's success. However, in general, the athlete's speed and quickness measurements are still done manually with a stopwatch, and then the results are written down in a book or paper. This study proposes a Quickness Aspect Talent Identification (QATI) system that uses Internet of Things (IoT) technology to identify athlete talent for quickness. This system is designed to obtain athlete quickness data by monitoring their movements through sensors placed at predetermined points and connected to a database via the Internet. The QATI system can help identify athletes that traditional talent identification methods might miss. The QATI system comprises four nodes, one controller, and a website-based information system. The node serves as an athlete's goal point, which can be placed at four points. The controller will set the on–off combination on the node with an adjustable duration or run with a random algorithm. NodeMCU sends time measurement results to a Wifi or Bluetooth network server. The measurement results from each sensor on the node will appear on the website that has been created. The QATI application is proven to perform several configurations such as training mode, delay between devices, duration of on, the distance the device detects, and viewing indicators of the number of connected sensors. The QATI system is an innovative approach to talent identification that utilizes IoT technology to identify athletes for quickness.

Keywords Athlete · Quickness · Speed · IoT · Qati

R. Tulloh (✉) · J. M. L. Gaol · A. Mulyana
School of Applied Science, Telkom University, Jawa Barat, Indonesia
e-mail: rohmatth@telkomuniversity.ac.id

D. Rimasa
Faculty of Sports and Health Education, Indonesia University of Education, Bandung, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_25

1 Introduction

After knowing the athlete's skill, it usually continues to identify talent. Talent identification is a significant factor in the search for qualified athletes. Essential aspects in many sports are speed, agility, and quickness aspects. In recent years, technology in sports, especially in monitoring athletes' physical development and talent, has been widely used. Internet of Things (IoT) technology is one of the trends in sports development. In applying IoT technology, the identification of athlete talent becomes more accessible. In the badminton branch, for example, a study conducted by Rusdiana [1], made measuring instruments for training reflexes and training the speed of badminton athletes. This tool uses an application that has been made, called Agility LED Control, from the Warriornux application system. This system is designed to identify the agility aspect of badminton athletes using infrared sensors to detect objects within the sensor's range. Another study by Wang et al. utilized IoT technology to make rackets that can retrieve data from basic badminton movements and collect data via the cloud [2]. Several other studies on similar sports, such as those by Xiong [3], or utilizing motion signals based on big data to determine the player's balance [4]. Other research that utilizes IoT to identify talent, such as research in the field of basketball [5, 6], tennis [7], football [8, 9], sport training [10, 11], sport community [12] or the use of sports training assistants for ice and snow sports training [13]. In another study, Liu made an IoT-based talent identification model for rowing [14]. Research conducted by Almasi also aims to determine talent identification in athletic athletes by utilizing human movement using a camera [15]. Nasir et al. Research in athletics using IoT was also carried out using data obtained from sensors placed on patients [16].

Many characteristics of an athlete may be assessed with internet of things technology, including speed, agility, quickness, endurance, strength, accuracy, and flexibility. Several studies have focused on evaluating physical features of athletes while they workout [17–19], exercise performance evaluation [20], body index monitoring [21] or sport health monitoring [22]. This study focuses on measuring quickness because not much research leads to measuring these aspects. The Internet of Things is used for training improvement using the cone drill technique, commonly used in training for athletes. This study also made an application that can be used to manage, view, and store training results in data. The trials were carried out on fencing athletes. Fencing is a sport that focuses on skill techniques such as cutting, stabbing, or parrying an opponent's weapons by utilizing hand agility [23]. In fencing, reflexes are an essential factor. This movement is a spontaneous movement that is carried out unconsciously when getting a stimulus (stimulus). The factor determining victory in this sport is how many touches the opponent's body area. So concentration is needed to touch the opponent and reflex movements to avoid the opponent's attack [24].

2 Materials and Method

2.1 Quickness Training

Quickness Aspect Talent Identification is a way of knowing an individual's talent or potential regarding quick reaction or agility. It is essential to consider that agility or fast reaction in sports is one of the crucial things a professional athlete must have besides skill, speed, and agility [25]. Aspects of agility and quick response are very influential in high-intensity sports, such as football, basketball, badminton, and fencing. The cone drills method is one of the training menus to increase the quickness aspect. The cone drills method is a training method that athletes commonly use to train speed, quick reactions, and concentration. As in Fig. 1. The form of the exercise involves cones or objects as markers placed in a regular pattern. Patterns can be straight lines, triangular patterns, or quadrilaterals. The athlete will be asked to run past the cone and return to the starting position.

In the early stages of the design, a block diagram of the Talent Identification Aspect Quickness tool was made, as shown in Fig. 2. The device consists of 4 nodes, one controller, and a website-based application for the user interface on a smartphone or laptop. One node represents one cone in the exercise. One node consists of an ESP8266 MCU node, one HC-SR04 Ultrasonic sensor, one buzzer to generate sound, and three red, green, and yellow LEDs. Lithium 18,650 batteries are used for power supply. TP40560 is used for charging Lithium 18,650 batteries for NodeMCU purposes. The XC6206 regulator is used to regulate the voltage on the NodeMCU. The four nodes will be controlled by a controller consisting of a NodeMCU ESP8266 and a battery. The controller is a liaison between the device and website-based applications using the internet network. On the controller, several features can be adjusted, including the delay between nodes, the duration of time each node is on, and the minimum and maximum distances of the ultrasonic sensor when detecting the athlete's hand (vertical). This tool has two modes, namely sequential mode, and random mode. In sequential mode, the lights on the nodes will light up sequentially, starting from nodes 1, 2, 3, and 4. Then repeat from number 1 until it is finished. In

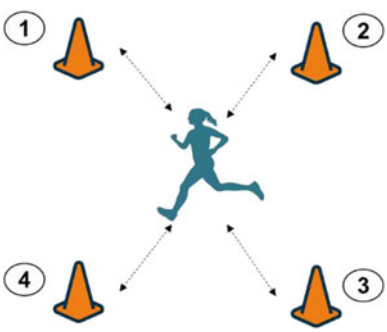


Fig. 1 A drill training scheme for quikness exercise

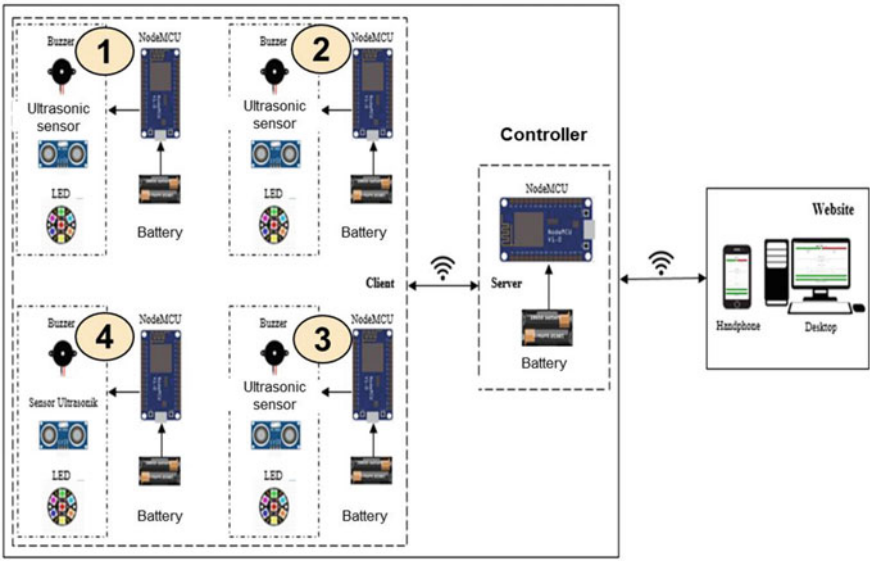


Fig. 2 Block diagram system

random mode, the sequence of nodes that will light up cannot be predicted. Random mode is designed to train athletes' concentration.

The athlete's task is to stand in the middle between the nodes, then run towards the active node (light changes from green to yellow) to deactivate the node (light changes from yellow to green) using hand movements until the ultrasonic sensor detects movement. After success, the athlete must return to all positions, and so on, until the program is complete. If the athlete is late in deactivating the node, the green light will change to red briefly before turning back to green). Data from the detection results from each node in the form of success points, and failure results, will be sent to the controller to be displayed in the application.

The workings of the Talent Quickness system can be explained regarding how it works, starting from obtaining configuration data in the form of delays between tools, the duration of time when the tool is on, and the minimum and maximum distance of the sensor when it detects. The next step is selecting the test pattern applied to each node. Sequential pattern or random pattern. The results of the configuration data that has been created are then stored and inputted through the website application and sent to the controller for further reading by each node. The features of the application can be seen in Table 1.

A Use Case Diagram illustrates the user interaction relationship between the system and the user. Figure 3 illustrates the use case diagram input values of measurement parameters that can be made. The user can control the tool by using the Start and Stop buttons and can save the configuration results via the Save Configuration button, consisting of pattern selection, tool delay, tool duration, and tool distance. Users can also see the number of connected clients via the Connected Devices button.

Table 1 Features in the application

No	Button	Output
1	Start	Node on, timer on
2	Stop	Node off, timer off
3	Reset	Reset timer, reset poin
4	Select a training mode	Sequential mode or random mode
5	Set Interval LED	How long the lights on the node are on
6	Set distance	Ultrasonic sensor max. detection distance
7	Save Configuration	Save Configuration
8	Connected node	Displays the number of connected nodes

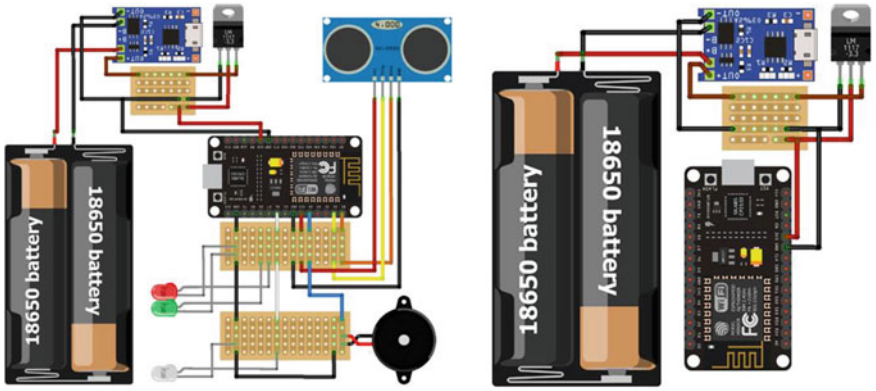


Fig. 3 Wiring diagram for node (left) and controller (right)

2.2 Client-Control System

NodeMCU, HC-SR04 Ultrasonic sensor, buzzer, and LED are used for client design. Ultrasonic sensors capture athletes' movements, while LEDs and buzzers are tool indicators. Lithium 18,650 batteries are used for power supply. TP40560 is used for charging Lithium 18,650 batteries for NodeMCU purposes. The XC6206 regulator is used to regulate the voltage on the NodeMCU.

On the NodeMCU controller a local network is created. The local network functions to connect between nodes and controllers and to open websites by entering the IP that has been created by the server. Component design for servers only uses NodeMCU, 18,650 Lithium battery, TP40560, and XC6206 regulator. A display of the schematic circuit on the server side can be seen in Fig. 3.

2.3 Node and Controller Integration (Client–server)

On the application function from the main page, there is a timer that functions to display the time from the test results, the start button to turn on the tool, and the stop button to stop the tool. Information about successful points, failed points, and the average speed obtained is provided. In the user interface created, two configuration buttons go to the configuration page and monitoring buttons. For the detailed function of each button can be seen in Table 1.

3 Result and Discussion

3.1 Hardware Impementation

The implementation results are shown in Fig. 4. Four nodes are designed using a transparent casing with 17 cm × 12 × cm × 8 cm dimensions. The transparent casing was chosen to make it easier to see the tool's performance and the lights on. Look at Figure 4. The node's condition is connected to the controller and website application.

The next test is a QATI trial on nine athletes willing to try the device. As in Fig. 5. Each node is placed at each corner to form a square with a distance of 5 m × 5 m in length and width. Then the athlete stands in the center, waiting for a signal to start.

The first experiment was carried out in regular mode 12 times. The node was lit with an interval of 10 s. From the measurement results displayed on the website the data is as shown in Table 2. This device can show quick-ness aspect test results from the measurement results through the number of successful points, the number

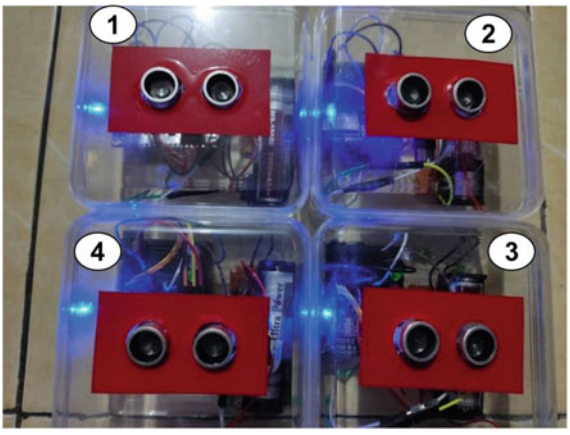


Fig. 4 Four nodes of QATI

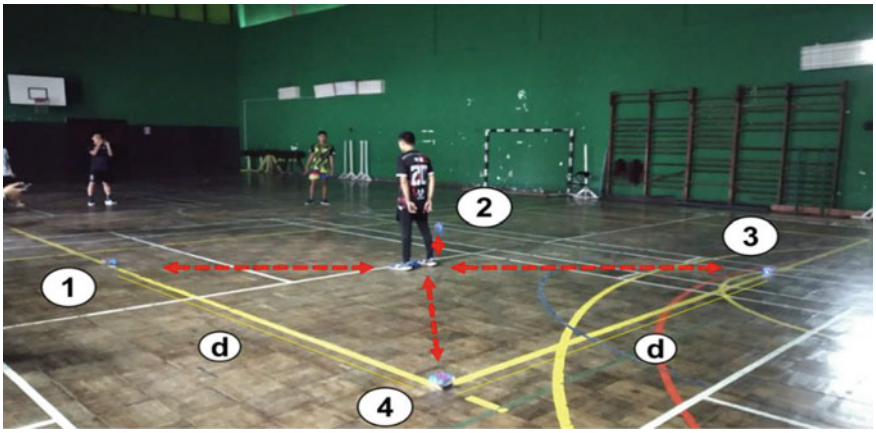


Fig. 5 Four nodes of QATI with distance d × d

of failed points, and the time taken. From the measurement results, it can also be seen who is the best and who needs more improvement.

Tests were also carried out to determine the maximum distance between the node and the controller and see the resulting delay. The test is performed by placing four nodes to form an equal rectangle. The test results can be seen in Table 3.

In Table 2, it can be seen that when nodes are placed at a distance of under 10 m × 10 m, all nodes are still detected and produce delays that are still in a suitable category, ranging from 150–300 ms whereas for excellent data packet transfer delays must be below the figure of 150 ms [26]. However, the delay will increase when the distance exceeds 10 × 10 m. The maximum distance where the node is still detected is at 24 m.

Table 2 Quickness aspect measurement results in athletes

Atlete Number	Succesful Points	Failed Points	Time
1	10	2	2 min 07 s
2	11	1	2 min 13 s
3	11	1	2 min 02 s
4	12	0	1 min 58 s
5	11	1	1 min 58 s
6	11	1	2 min 7 s
7	10	2	1 min 58 s
8	10	2	2 min 9 s
9	10	2	2 min 10 s

Table 3 Testing delay based on distance

No	Distance (d × d m)	Node 1 (s)	Node 2 (s)	Node 3 (s)	Node 4 (s)
1	3 × 3	0.17	0.21	0.20	0.18
2	5 × 5	0.24	0.23	0.23	0.22
3	10 × 10	0.30	0.31	0.29	0.29
4	15 × 15	0.49	0.52	0.53	0.54
5	24 × 24	1.20	1.10	1.20	1.20
6	25 × 25	–	–	–	–
7	30 × 30	–	–	–	–

3.2 Website Impementation

The features of the website are shown in Fig. 6. In the initial conditions, it appears that all features can be displayed, such as test time, training mode, success points, failed points, the duration of the device is on, the distance the ultrasonic sensor detects, saves configurations and connections are connected to each node.

In Fig. 7. The website application can display the results of quickness training. The test results show that the application can display the length of time the test was carried out, the number of successful and failed points. The time it takes for the

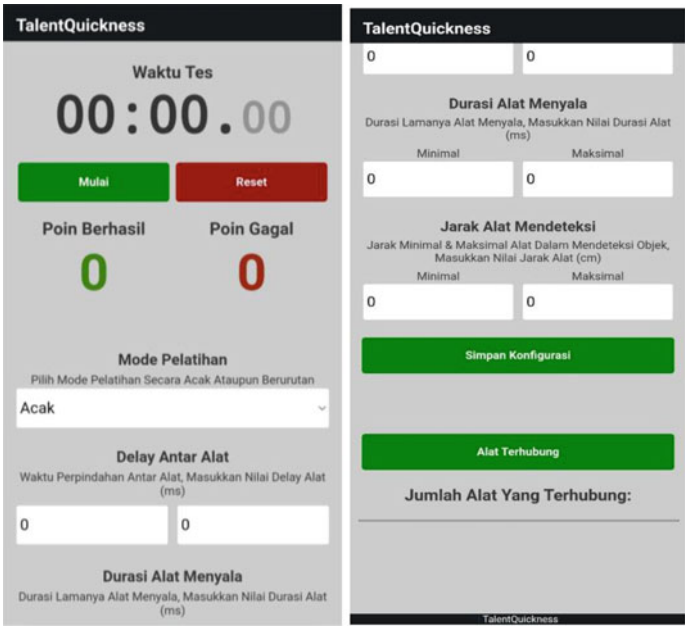


Fig. 6 User interface web application

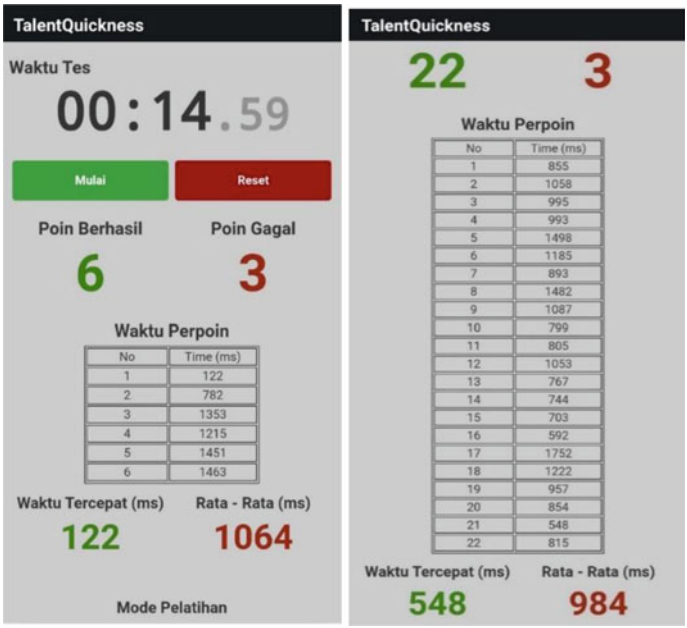


Fig. 7 Result testing on web application

athlete to deactivate each node. As well as the fastest time and the average time generated by athletes.

4 Conclusion

Based on the results of the design and testing of the tools that have been carried out, the following conclusions are obtained. Based on the tests carried out, the Internet of Things-based quickness aspect talent identification (QATI) tool has succeeded in measuring and displaying the results of the quickness level in cone drills for athletes. The Settings feature on the website-based application can configure training mode, delay between tools, duration of time for each tool to turn on, and the minimum and maximum distances for ultrasonic sensors to detect objects in athletes' hands. The maximum distance between the controller and the node for the client is a maximum of 24 m. For longer distances, external antennas can be added so that the signals of each node can be detected. In future research, we aim to add features to different training methods so that the plans for measuring the quickness aspect are more varied.

References

1. Rusdiana A (2021) Development of agility, coordination, and reaction time training device with infrared sensor and WiFi module Arduino in badminton. Songklanakarin J Sci Technol 2(43)
2. Wang Y et al (2018) IoT for next-generation racket sports training. IEEE Internet Things J 5(6). https://doi.org/10.1109/JIOT.2018.2837347
3. Xiong S (2022) Intelligent strategy of internet of things computing in badminton sports activities. Hindawi Wirel Commun Mob Comput 2022:9
4. Jitpattanaku SM, Anuchit PJNH (2022) Badminton activity recognition and player assessment based on motion signals using deep residual network. 2022 IEEE 13th international conference software engineering service science
5. Okolyi M, Elshakankiri K (2020) Internet of things in the game of basketball. In IoT as a Service. In: 6th EAI international conference, IoTaaS 2020, pp 421–435
6. Chen W (2021) Internet of things and edge computing model optimization in physical training status and countermeasures in college sports basketball optional course. Comput Intell Neurosci 1(11)
7. Hu R (2023) IoT-based analysis of tennis player's serving behavior using image processing. Soft Comput 27(19):14413–14429. https://doi.org/10.1007/s00500-023-09031-w
8. Wang J, Liu B (2023) Analyzing the feature extraction of football player's offense action using machine vision, big data, and internet of things. Soft Comput 27(15):10905–10920. https://doi.org/10.1007/s00500-023-08735-3
9. Singh MK, Singh AK, Singh P, Rishi OP (2023) Artificial intelligence enabled IOT system for football identification in a football match. In: Communications in computer and information science, vol 1782 CCIS, pp 460–472. https://doi.org/10.1007/978-3-031-35644-5_37
10. Zhang Y, Duan W, Villanueva LE, Chen S (2023) Transforming sports training through the integration of internet technology and artificial intelligence. Soft Comput 27(20):15409–15423. https://doi.org/10.1007/s00500-023-08960-w
11. Chen M et al (2023) Phase-based quantification of sports performance metrics using a smart IoT sensor. IEEE Internet Things J 10(18):15900–15911. https://doi.org/10.1109/JIOT.2023.3266351
12. Yin Z, Li Z, Li H (2023) Application of internet of things data processing based on machine learning in community sports detection. Prev Med (Baltim) 173. https://doi.org/10.1016/j.ypmed.2023.107603
13. Zhang J, Sun P (2022) Application practice analysis of ice and snow sports training assistance system based on internet of things. Wirel Commun Mob Comput
14. Liu J, Chen S, Wang C (2020) An intelligent identification model for the selection of elite rowers by incorporating internet-of-things technology. IEEE Access 8:31234–31244. https://doi.org/10.1109/ACCESS.2020.2973418
15. Ghaeinian SA(2020) Investigating the application of human motion recognition for athletics talent identification using the head-mounted camera. In: Fifth International conference on inventive computation technologies (ICICT-2020), pp 54–59
16. Nasir JM, Mamun M, Rahman MM, Mahmud MIA, Sabbir Bin MH (2022) MHfit: mobile health data for predicting athletics fitness using machine learning models. https://doi.org/10.1109/ISMODE56940.2022.10180967
17. Li W (2023) Application of IoT-enabled computing technology for designing sports technical action characteristic model. Soft Comput 27(17):12807–12824. https://doi.org/10.1007/s00500-023-08966-4
18. Zhao J, Zhao Y, Wang H (2023) Simulation of sports training recognition system based on internet of things video behavior analysis. Prev Med (Baltim) 173. https://doi.org/10.1016/j.ypmed.2023.107611
19. Ning C, Li M, Ge L (2023) The preventive effect of PNF stretching exercise on sports injuries in physical education based on IoT data monitoring. Prev Med (Baltim) 173. https://doi.org/10.1016/j.ypmed.2023.107591

20. Sun S, Peng T, Huang H, Wang Y, Zhang X, Zhou Y (2023) IoT motion tracking system for workout performance evaluation: a case study on dumbbell. IEEE Trans Consum Electron, 1. https://doi.org/10.1109/TCE.2023.3320183
21. Gong H, Si Y (2023) Simulation of a sports body index monitoring system based on internet of things technology. Lect Notes Data Eng Commun Technol 167:273–280. https://doi.org/10.1007/978-981-99-0880-6_30
22. Zeng W, Martínez OS, Crespo RG (2023) Energy harvesting IoT devices for sports person health monitoring. J Ambient Intell Humaniz Comput 14(4):3727–3738. https://doi.org/10.1007/s12652-021-03498-x
23. Harfoush A, Hossam M (2020) Modelling of a robot-arm for training in fencing sport. Int J Intell Robot Appl 4:109–121. https://doi.org/10.1007/s41315-020-00116-5
24. Doron J (2021) Dealing with elite sport competition demands: an exploration of the dynamic relationships between stress appraisal, coping, emotion, and performance during fencing matches. Cogn Emot, 1365–1381. https://doi.org/10.1080/02699931.2021.1960800
25. Rimasa D (2021) Anthropometric, biomotor, and physiological factors for indonesian fencing athletes talent identification. J Terap Ilmu Keolahragaan 6(1):6–13. https://doi.org/10.17509/jtikor.v6i1.31191
26. Tulloh R, Amri Ginting JG, Mulyana A, Lutfi M (2020) Performance comparison of file transfer protocol service between link state and distance vector routing protocol in software defined network. In: IOP conference series: materials science and engineering, vol 982, no 1. https://doi.org/10.1088/1757-899X/982/1/012026

Sentiment Analysis on Service Quality of an Online Healthcare Mobile Platform Using VADER and Roberta Pretrained Model



Fairuz Iqbal Maulana, Puput Dani Prasetyo Adi, Dian Lestari, Agung Purnomo, and Daniel Anando Wangean

Abstract The development of mobile technology has penetrated into the health sector. With the development of this technology, information about health can improve the quality of health independently. The integration of information technology into the healthcare sector, namely the advancement of mobile-based health services, has significantly revolutionised the accessibility of healthcare across various regions in Indonesia. Halodoc is the leading digital health company in Indonesia and has substantially changed the axis of health services in Indonesia by providing health information that is easy to understand, credible, and accessible to everyone. This research will analyze Halodoc's service quality based on customer reviews from Google Play Store, which will be analyzed using sentiment analysis. Specifically, this research uses VADER (Valence Aware Dictionary and Sentiment Reasoner) and Roberta Pretrained Model to analyze the sentiment of user reviews on Halodoc mobile application. The results of the sentiment analysis can provide insight into users' overall satisfaction with the platform's service quality. Results show that 84% users leave reviews with Neutral Sentiment Analysis, 13% with positive comments, and the smallest number of Sentiment Analysis is negative with 3%, in this application. We create correlation matrix to see correlation coefficients between sets of variables and the wordcloud. This research contributes to the growing body of research

F. I. Maulana (✉) · D. A. Wangean
Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta 11480, Indonesia
e-mail: fairuz.maulana@binus.edu

P. D. P. Adi
Telecommunication Research Center, National Research and Innovation Agency, Jakarta 11480, Indonesia

D. Lestari
Faculty of Medicine, Universitas Airlangga, Surabaya 60132, Indonesia

A. Purnomo
Entrepreneurship Department, BINUS Business School Undergraduate Program, Bina Nusantara University, Jakarta 11480, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_26

on sentiment analysis in the healthcare industry and can inform the development of strategies to improve the quality of online healthcare services.

Keywords Machines learning · Sentiment analysis · Medical services · Internet · VADER · Information technology

1 Introduction

The rapid development of mobile and internet technologies has revolutionized the healthcare industry [1], with numerous digital platforms emerging to offer healthcare services at the click of a button [2]. In particular, online healthcare mobile platforms [3] have become increasingly popular due to their convenience and the high demand for remote medical services, especially in the wake of the COVID-19 pandemic [4, 5]. These platforms facilitate communication between healthcare professionals and patients, prescription management, appointment scheduling, and much more. As these platforms become more prevalent, evaluating and maintaining the quality of the services they offer is paramount. In this context, the current research aims to analyze the sentiment of user reviews for an online healthcare mobile platform using the VADER (Valence Aware Dictionary and sEntiment Reasoner) [6] and RoBERTa (Robustly Optimized BERT Pretraining Approach) [7] models.

The service quality of an online healthcare mobile platform is multi-faceted, encompassing aspects like ease of use, security, responsiveness, reliability, and the quality of healthcare provided. Customer reviews for such platforms are rich, untapped sources of data that offer insightful information regarding user satisfaction and areas for improvement. However, analyzing this large and unstructured data requires robust Natural Language Processing (NLP) techniques [8]. Sentiment analysis, or opinion mining, is one of the most effective NLP methods used to interpret and understand the sentiments [9], opinions [10], or emotions expressed in a piece of text [11].

VADER is a sentiment analysis tool that use a vocabulary and rule-based approach to effectively identify sentiment conveyed in social media. Its applicability extends to the concise and casual language frequently encountered in user reviews [12]. It is pre-built with a lexicon that assigns intensity scores to each word based on its sentiment, which makes it very effective at handling a wide range of sentiments. One of the strengths of VADER is its ability to understand the context of words and recognize sentiment in text with minimal preprocessing [13].

RoBERTa, on the other hand, is a more advanced, deep learning-based model that has been pretrained on a large corpus of text data [7]. It is built on the BERT (Bidirectional Encoder Representations from Transformers) model [14], with improvements in training methodology and data. RoBERTa understands the context in which a word is used and can interpret complex sentence structures and semantics. Due to its deep learning nature, it requires significant computing resources but often results in highly accurate sentiment classification.

Previous research in the area of sentiment analysis has been extensive, with numerous studies employing various techniques for different domains. For instance, studies have employed sentiment analysis to evaluate customer reviews in retail [15], hospitality [16], and online services [17], illustrating the effectiveness of these methods in extracting actionable insights from consumer-generated content.

In healthcare, previous research has employed sentiment analysis to assess patient satisfaction from hospital feedback forms and social media comments [18], demonstrating the applicability of these methods in this domain. One notable study utilized the VADER model to analyze patient reviews of healthcare providers, highlighting the ability of this model to accurately classify reviews and identify areas for improvement.

While prior studies have provided significant insights into the application of sentiment analysis in various domains, there is limited research focusing specifically on online healthcare mobile platforms. Moreover, the comparative analysis of a lexicon-based model like VADER and a deep learning-based model like RoBERTa in this specific context represents a novel contribution. This research aims to fill this gap by applying these two different approaches to a single dataset of user reviews for an online healthcare mobile platform, thereby offering a comprehensive and comparative analysis.

Given this background, the contribution of this study:

a. Provides insights into the sentiment of users towards the service quality of an online healthcare mobile platform.
b. The results of the study can be used by the platform to improve its service quality and user experience.
c. It can be used by online healthcare mobile platforms to improve their service quality and user experience.

2 Materials and Methods

2.1 System Overview

Researchers divide the system into two, namely the system to retrieve data from the source and the system to process the data that has been obtained. The system retrieves data through the Indonesian Google Play Store with a total of the last 2000 data. Data taken on July 28, 2023. Data is retrieved using Python programming language by calling google-play-scraper library. The 2000 data collected is extracted into CSV format files. The csv file is then processed on the second system through Google Colab for sentiment analysis using several techniques such as VADER—Bag of words approach, and Roberta Pretrained Model [19].

Table 1 Result of the cleaning process

Content before cleaning	Content after cleaning
Terimakasih untuk aplikasi ini sangat membantu dimana pun dan kapanpun untuk berkonsultasi oleh para ahlinya. Sehat selalu dokter □	Terimakasih untuk aplikasi ini sangat membantu dimana pun dan kapanpun untuk berkonsultasi oleh para ahlinya. Sehat selalu dokter

2.2 Data Collection and Cleaning

Data gathering is conducted via utilising the Google Play Store platform in conjunction with Google Colab, which interfaces with the google-play-scraper library. The aforementioned procedure is implemented using the Python programming language and incorporates keywords derived from links originating from programmes available on the Google Play Store. The data searched in Indonesian and amounted to the last 2000 data. The results of crawling the data are exported into CSV files or Comma Separated Values.

The next step is cleaning to remove noise and content that is not related to data from the Google Play Store. This process is carried out by visual inspection. The cleaning procedure also involves the removal of URLs, punctuation, and number marks. Furthermore, the cleaning procedure involves the elimination of user mentions denoted by the "@" symbol and the removal of duplicate tweets. The outcomes of the cleansing procedure are observable in Table 1.

2.3 Data Preprocessing

After the data is cleaned, the next process is for researchers to do quick EDA (Exploratory Data Analysis) to analyze review stars from 2000 data of various users. Then continued with basic NLTK, followed by VADER and Roberta Pretrained Model to analyze the sentiment of user reviews on Halodoc mobile application. Sentiment analysis can also be done by measuring polarity and subjectivity. Polarity refers to the strength of an opinion which is positive or negative. Subjectivity refers to the extent to which a person is personally involved in an object. The most important thing here is the personal relationship and experience of the individual with the object, which may differ from the point of view of others (Fig. 1).

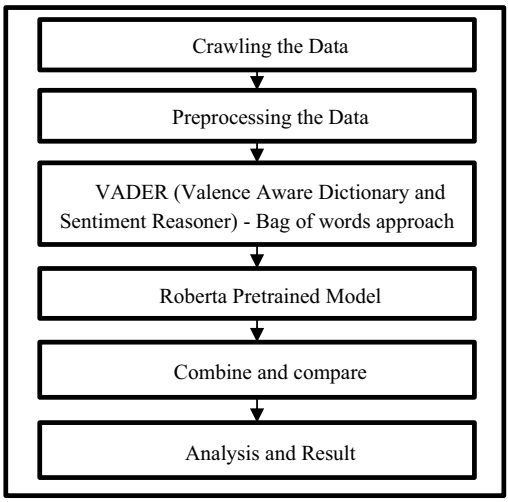


Fig. 1 Flow process sentiment analysis on service quality of an online healthcare mobile platform using vader and roberta pretrained model

3 Result and Analysis

3.1 Quick EDA

The preliminary examination of data, commonly referred to as Exploratory Data Analysis (EDA), has significant importance within the context of any data analysis endeavour, including sentiment analysis. It involves examining and summarizing the main characteristics of the data to gain insights and identify patterns. In the context of sentiment analysis on the service quality of an online healthcare mobile platform, EDA can help us understand the distribution of sentiment scores, the most common words and phrases used in positive and negative reviews, and the relationship between sentiment and other variables such as time of day or user demographics. One quick way to perform EDA on the sentiment data is to create visualizations such as histograms, word clouds, and scatterplots. Researchers visualized Review Stars in Google Play Store from 2000 data obtained and divided it into stars 1–5 (Fig. 2).

Nevertheless, it is crucial to acknowledge that relying just on star ratings may not offer a comprehensive understanding of mood, as users have the tendency to write comments that may not correspond with their assigned star rating. Therefore, it is important to also analyze the text of the reviews using sentiment analysis tools such as VADER and Roberta Pretrained Model to gain a more comprehensive understanding of the sentiment towards the platform.



Fig. 2 Count of reviews by stars

Table 2 Polarity score on a sentence

Content	Negative score	Neutral score	Positive score
Alhamdulillah sangat membantu sekali disaat urgent. Tapi waktu itu pernah coba mau call, udah bayar tapi lama ga di angkat, akhirnya saya batalkan, dan dana di kembalikan. Terima kasih halo doc	0.0	0.943	0.057

3.2 VADER Sentiment

The VADER tool is a sentiment analysis tool that is extensively utilised in the field of natural language processing. It is designed to analyze the sentiment of text by assigning a positive, negative, or neutral score to each word based on its semantic orientation. VADER uses a combination of rule-based heuristics and machine learning techniques to accurately identify the sentiment of text, even in cases where the sentiment is ambiguous or sarcastic. The neg/neu/pos ratings of the text will be obtained using NLTK's SentimentIntensityAnalyzer. The methodology employed in this study adopts a "bag of words" approach. In the process of text analysis, the elimination of stop words is undertaken, followed by the assignment of scores to individual words, which are subsequently aggregated to get a cumulative score.

In Fig. 3, it shows the percentage of result who had a positive, neutral, or negative sentiment towards the platform. In positive sentiment, a 5-star rating indicates higher graphics or more positive reviews. In neutral sentiment, Star ratings 1 to 5 have the same number. Meanwhile, for negative sentiment, 1 and 2-star ratings are enough to indicate user reviews that are less than satisfied with the application.

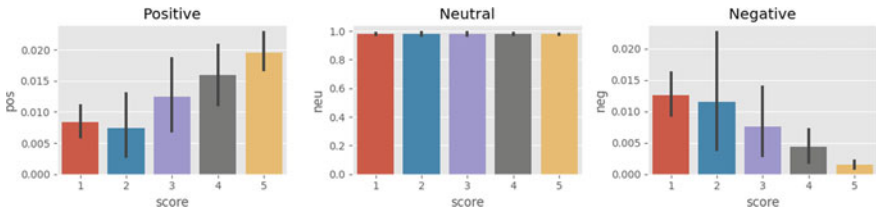


Fig. 3 Plot VADER result

From the implementation of VADER, researchers can visualize the score of each positive, neutral, or negative sentiment (Fig. 3). Then, researcher try to find polarity score from some sentence (Table 2).

3.3 Roberta Pretrained Model

The Roberta Pretrained Model is a very advanced language model that is extensively utilised in many natural languages processing applications, including sentiment analysis. The model in question is a variation of the BERT (Bidirectional Encoder Representations from Transformers) architecture, which undergoes pretraining on a substantial corpus of textual data in order to acquire knowledge about the contextual associations of words. Roberta uses a masked language modeling objective to learn the representations of words in context, which allows it to capture complex linguistic phenomena such as negation, sarcasm, and ambiguity. To use Roberta for sentiment analysis on the service quality of an online healthcare mobile platform, we use library transformer then researcher import Auto Tokenizer, Auto Model for Sequence Classification, and softmax. In Table 3, show the Roberta score on a sentence.

Table 3 Roberta score on a sentence show the high score in Roberta Neutral

Content	Roberta negative	Roberta neutral	Roberta positive
Sangat sangat dan sangat, membantu sya ketika kurang enak badan. Dalam konsultasi sampai Obat	0.2520102	0.7119708	0.03601897

3.4 *Compare Score Between Models (VADER and Roberta Pretrained)*

The comparison of sentiment scores between the VADER and Roberta Pretrained Model might yield useful insights regarding the performance and efficacy of these two sentiment analysis methodologies.

A method for comparing the scores involves the computation of the correlation between the emotion ratings produced by each model. This can help us determine the level of agreement between the two models in terms of sentiment classification. A high correlation would indicate that both models are consistent in their sentiment predictions, while a low correlation may suggest differences in their interpretations of sentiment. By analyzing the distribution of sentiment scores from each model, we can gain insights into the overall sentiment distribution and identify any discrepancies or differences in sentiment classification (Fig. 4).

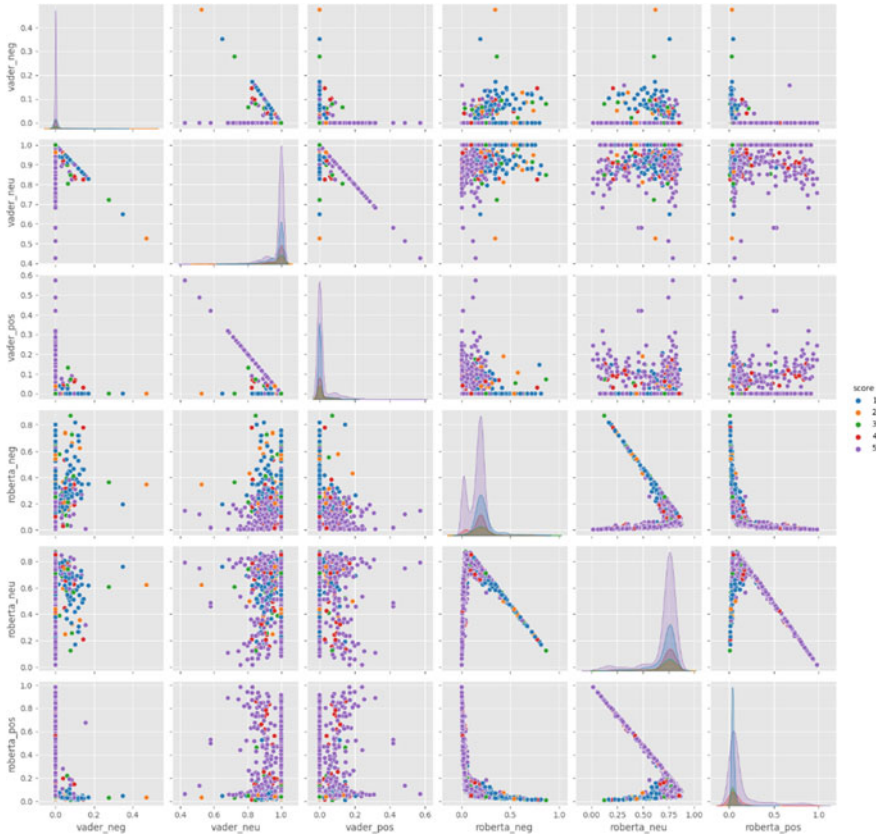


Fig. 4 Compare score between VADER and Roberta Pretrained Models

Table 4 Roberta score on a sentence

Neutral	Negative	Positive
84%	3%	13%



Fig. 5 Correlation heatmap for sentiment analysis on service quality of an online healthcare mobile platform

Furthermore, it is important to consider the strengths and limitations of each model. VADER is specifically designed for sentiment analysis and has been widely used in various domains. It is known for its simplicity and effectiveness in capturing sentiment in text. On the other hand, the Roberta Pretrained Model is a more advanced language model that can capture complex linguistic phenomena and context. It may provide more nuanced sentiment predictions but may also require more computational resources and expertise for fine-tuning and implementation.

3.5 Sentiment Analysis

From 2000 data that has been obtained when processed, the percentage value of sentiment analysis is obtained as in Table 3. Researchers also tried to visualize correlation heatmap (Fig. 5) and wordcloud (Fig. 6 and Table 4).

4 Discussion

From the results and analysis, it can be discussed as follows:

a. Interpretation of the finding:

   The study analyzed user reviews of Halodoc, an online healthcare mobile platform, using sentiment analysis techniques. The results showed that the majority of users left neutral reviews, while a small proportion of users left positive or



Fig. 6 Wordcloud from 2000 data for sentiment analysis on service quality of an online healthcare mobile platform

negative reviews. The study also developed a correlation matrix that identified features or services that significantly impacted user satisfaction. The word cloud visualization highlighted specific keywords or phrases frequently associated with users' sentiments, providing a clear guide for Halodoc to understand what matters most to its users. Overall, the study provides actionable insights for Halodoc to improve its service quality and solidify its position as a transformative player in Indonesia's healthcare sector.

b. Comparison to other studies:

Previous research has employed sentiment analysis to evaluate customer reviews in various domains [20], including retail, hospitality [16], and online services [17]. In healthcare, prior studies have used sentiment analysis to assess patient satisfaction from hospital feedback forms and social media comments. However, there is limited research focusing specifically on online healthcare mobile platforms. This study contributes to the growing body of research on sentiment analysis in the healthcare industry. It can inform the development of strategies to improve the quality of online healthcare services.

c. Weaknesses or limitations:

One limitation of the study is that it only analyzed user reviews from the Google Play Store, which may not be representative of all users of Halodoc. Additionally, the study did not consider other factors that may impact user satisfaction, such as the quality of healthcare provided or the responsiveness of healthcare professionals. Finally, the study did not compare the effectiveness of VADER and RoBERTa in analyzing user sentiment, which could be a valuable area for future research.

d. Implication to society:

The study's findings have important implications for the healthcare industry, particularly in the context of the COVID-19 pandemic, where remote medical services have become increasingly important. By analyzing user reviews, healthcare providers can gain valuable insights into user satisfaction and areas for improvement. This can inform the development of strategies to improve the quality of online healthcare services, making healthcare more accessible and convenient for people in Indonesia and beyond.

5 Conclusion

This study utilized VADER and RoBERTa Pretrained Models to conduct a comprehensive sentiment analysis of user reviews for Halodoc, Indonesia's leading digital health company, from the Google Play Store. The findings reveal a predominantly neutral sentiment in user reviews, with 84% of the users expressing neither extreme satisfaction nor dissatisfaction. Notably, 13% of the reviews were positive, highlighting features like credible information and accessibility, which users find beneficial. In stark contrast, a minimal proportion, 3%, of the reviews were negative, indicating minor areas that need attention, possibly relating to app functionalities or specific services offered through the platform.

The correlation matrix developed in this research provides a detailed and quantifiable relationship between sets of variables extracted from the user reviews, which offers Halodoc actionable insights into features or services that significantly impact user satisfaction. The word cloud visualization, a standout aspect of this study, amplifies specific keywords or phrases that are frequently associated with users' sentiments, providing a clear guide for Halodoc to understand what matters most to its users. These insights are invaluable, directing Halodoc towards data-driven refinements that can further enhance its service quality and solidify its position as a transformative player in Indonesia's healthcare sector.

References

1. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, Wang Y, Dong Q, Shen H, Wang Y (2017) Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2:230–243. https://doi.org/10.1136/svn-2017-000101
2. Marcos-Pablos S, Juanes Méndez JA, Walters ML (2020) State-of-the-art technologies at the service of medical training and practice to foster digital health ecosystems. In: Eighth international conference on technological ecosystems for enhancing multiculturality. ACM, New York, NY, USA, pp 410–413. https://doi.org/10.1145/3434780.3436700
3. Xu Y, Jiang Y, Li R, Gao H, Guo J, Liu Y, Hei L, Wang Y (2022) A healthcare-oriented mobile question-and-answering system for smart cities. Trans Emerg Telecommun Technol 33. https://doi.org/10.1002/ett.4012
4. Valdez D, ten Thij M, Bathina K, Rutter LA, Bollen J (2020) Social media insights into US mental health during the COVID-19 pandemic: longitudinal analysis of Twitter data. J Med Internet Res 22:e21418. https://doi.org/10.2196/21418

5. Maulana FI, Adi PDP, Lestari D, Purnomo A, Prihatin SY (2022) Twitter data sentiment analysis of COVID-19 vaccination using machine learning. In: 2022 5th international seminar on research of information technology and intelligent systems (ISRITI). IEEE, pp 582–587. https://doi.org/10.1109/ISRITI56927.2022.10053035
6. Hutto C, Gilbert E (2014) VADER: a parsimonious rule-based model for sentiment analysis of social media text. Proc Int AAAI Conf Web Soc Media 8:216–225
7. Yohannes HM, Amagasa T (2022) Named-entity recognition for a low-resource language using pre-trained language model. In: Proceedings of the 37th ACM/SIGAPP symposium on applied computing. ACM, New York, NY, USA, pp 837–844. https://doi.org/10.1145/3477314.3507066
8. Samant RM, Bachute MR, Gite S, Kotecha K (2022) Framework for deep learning-based language models using multi-task learning in natural language understanding: a systematic literature review and future directions. IEEE Access 10:17078–17097. https://doi.org/10.1109/ACCESS.2022.3149798
9. Khalid ET, Talal EB, Faraj MK, Yassin AA (2022) Sentiment analysis system for COVID-19 vaccinations using data of Twitter. Indones J Electr Eng Comput Sci 26:1156–1164. https://doi.org/10.11591/ijeecs.v26.i2.pp1156-1164
10. Prasetyo F, Putra E, Iqbal F, Muhammad N (2023) Twitter sentiment analysis about economic recession in indonesia. Bull Soc Informatics Theory Appl 7:1–7. https://doi.org/10.31763/businta.v7i1.592
11. Khan R, Khan R, Shrivastava P, Kapoor A, Tiwari A, Mittal A, Head (2020) Social media analysis with AI: sentiment analysis techniques for the analysis of Twitter Covid-19 data. Artic J Crit Rev 7:2765–2766
12. Chandrasekaran R, Mehta V, Valkunde T, Moustakas E (2020) Topics, trends, and sentiments of Tweets about the COVID-19 pandemic: temporal Infoveillance study. J Med Internet Res 22:e22624. https://doi.org/10.2196/22624
13. Mujahid M, Lee E, Rustam F, Washington PB, Ullah S, Reshi AA, Ashraf I (2021) Sentiment analysis and topic modeling on Tweets about online education during COVID-19. Appl Sci 11. https://doi.org/10.3390/app11188438
14. Acikalin UU, Bardak B, Kutlu M (2020) Turkish sentiment analysis using BERT. In: 2020 28th signal processing and communications applications conference SIU 2020—proceedings. https://doi.org/10.1109/SIU49456.2020.9302492
15. Sindhu C, Rajkakati D, Shelukar C (2021) Context-based sentiment analysis on amazon product customer feedback data. https://doi.org/10.1007/978-981-15-5329-5_48
16. Mehraliyev F, Chan ICC, Kirilenko AP (2022) Sentiment analysis in hospitality and tourism: a thematic and methodological review. Int J Contemp Hosp Manag 34:46–77. https://doi.org/10.1108/IJCHM-02-2021-0132
17. Pandesenda AI, Yana RR, Sukma EA, Yahya A, Widharto P, Hidayanto AN (2020) Sentiment analysis of service quality of online healthcare platform using fast large-margin. In: Proceedings of the 2nd international conference on informatics, multimedia, cyber, information systems ICIMCIS 2020, pp 121–125. https://doi.org/10.1109/ICIMCIS51567.2020.9354295
18. Ahmad M, Aftab S, Ali I (2017) Sentiment analysis of Tweets using SVM. Int J Comput Appl 177:25–29. https://doi.org/10.5120/ijca2017915758
19. ROB MULLA: Sentiment analysis Python ? [Youtube Tutorial]. https://www.kaggle.com/code/robikscube/sentiment-analysis-python-youtube-tutorial. Last accessed 01 April 2023
20. Jiang Y, Wang H, Yi T (2021) Evaluation of product reviews based on text sentiment analysis. In: ACM international conference proceeding series. https://doi.org/10.1145/3469213.3470379

Internet of Medical Things: A Bibliometric Analysis of Research Publications from 2018–2022

Fairuz Iqbal Maulana, Dian Lestari, Puput Dani Prasetyo Adi, Mohammad Nazir Arifin, and Agung Purnomo

Abstract The field of the Internet of Medical Things (IoMT) is experiencing significant growth as it involves the integration of medical devices and systems with the internet and various digital technologies. This study provides a bibliometric analysis of scholarly articles pertaining to the Internet of Medical Things (IoMT) during the time frame of 2018 to 2022. The analysis is conducted using data sourced from the Scopus database. A total of 1165 documents were identified and analyzed to gain insights into the current state of research in this field. The analysis includes trends in publication output, top authors and institutions, and key research topics. The results of the analysis indicate a steady increase in publication output on IoMT over the past five years, with the highest number of publications in 2022. The top authors in this field were Mohanty, S.P., Kougianos, E., and Pirbhulal, S., while the top institutions were King Saud University, Chinese Academy of Sciences, and University of North Texas. Co-citation analysis revealed that wearable devices, telemedicine, and data analytics were among the key research topics in IoMT. This bibliometric analysis offers useful insights into the present research landscape of the Internet of Medical Things (IoMT) spanning the years 2018 to 2022. The findings can be beneficial for researchers, practitioners, policymakers, and funding agencies in understanding the trends and priorities in this field. The top authors and institutions identified in this

F. I. Maulana (✉)
Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta 11480, Indonesia
e-mail: fairuz.maulana@binus.edu

D. Lestari
Faculty of Medicine, Universitas Airlangga, Surabaya 60132, Indonesia

P. D. P. Adi
Telecommunication Research Center, National Research and Innovation Agency, Jakarta 11480, Indonesia

M. N. Arifin
Informatics Engineering Study Program, University of Madura, Pamekasan, Indonesia

A. Purnomo
Entrepreneurship Department, BINUS Business School Undergraduate Program, Bina Nusantara University, Jakarta 11480, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_27

analysis can serve as valuable resources for collaboration and staying updated on the latest developments in IoMT.

Keywords Internet of medical things · Healthcare informatics · Bibliometric · VosViewer · Co-citation

1 Introduction

The field of the Internet of Medical Things (IoMT) is experiencing significant growth as it encompasses the integration of medical devices and systems with the internet and various digital technologies [1], which is a network of physical devices [2], vehicles, home appliances, and other items embedded with electronics, software, sensors, and connectivity [3–5]. The Internet of Medical Things (IoMT) [6] possesses the capacity to significantly transform the healthcare sector through its ability to facilitate remote patient monitoring, tailored treatment, and real-time data analysis [7]. With the increasing adoption of IoMT devices and systems, there is a need for research to understand the current state of IoMT and identify key research topics and trends [3, 8–10]. The IoT has been the subject of extensive research in recent years, with numerous publications exploring its applications, challenges, and opportunities [11].

The IoMT has emerged as a promising application of the IoT in the healthcare sector [12–15]. The IoMT can enable healthcare providers to monitor patients remotely, collect real-time data on their health status, and provide personalized treatment options. The IoMT can also improve the efficiency of healthcare delivery by reducing the need for in-person visits and enabling early detection of health issues [16, 17]. Hassan et al. [18] conducted a comprehensive survey of the IoT security, including the IoMT.

Despite the potential benefits of the IoMT, there are also significant challenges [19] that need to be addressed. One of the main challenges is the security [18, 20] and privacy of the data collected by the IoMT devices [18, 21–23]. The IoMT devices can collect sensitive health information, which needs to be protected from unauthorized access and use. To address this need, a bibliometric analysis of research publications on IoMT from 2018 to 2022 was conducted using data from the Scopus database [24]. The analysis included trends in publication output, top authors and institutions, and key research topics [25, 26]. A total of 1166 documents were identified and analyzed to gain insight into the current state of research in this area.

The analysis results show a steady increase in publication output in IoMT over the past five years, with the highest number of publications in 2022. The top authors in this field are Mohanty, S.P., Kougianos, E., and Pirbhulal, S., while the top institutions are King Saud University, Chinese Academy of Sciences, and University of North Texas. Citation analysis revealed that wearables, telemedicine, and data analytics are among the key research topics in IoMT.

This bibliometric analysis offers useful insights into the present condition of research on the Internet of Medical Things (IoMT) within the time frame of 2018–2022. The findings can be useful for researchers, practitioners, policymakers, and funding agencies in understanding trends and priorities in this field [18, 27, 28]. The leading authors and institutions identified in this analysis can be a valuable resource to collaborate with and keep abreast of the latest developments in IoMT.

Overall, the IoMT is a rapidly growing field with significant potential to transform healthcare [29–31]. The results of this bibliometric analysis highlight the need for continued research and collaboration in this area to fully realize the benefits of IoMT. By analyzing the trends, authors, and publications in the field, I hope to gain insights into the current state of research on the IoMT and identify areas for future research. The results of this analysis can be useful for researchers, healthcare providers, and policymakers who are interested in the IoMT and its potential applications in healthcare.

2 Methods and Materials

2.1 Data Source and Search Strategy

The primary aim of this study is to examine and evaluate the prevailing patterns and directions of research pertaining to the subject matter of Internet of Medical Things. Bibliometric analysis is employed because of its high precision in quantifying and assessing scholarly publications [32]. In addition, the acquisition of data is made more efficient through the application of the Scopus database. Scopus and Web of Science (WoS) are widely recognized as the primary and comprehensive sources for undertaking bibliometric analysis. However, Scopus possesses some notable advantages over WoS. Firstly, Scopus offers a more extensive coverage of content, en-compassing a wider range of scholarly materials. Additionally, Scopus is more accessible to the academic community, fostering a greater sense of openness. Furthermore, Scopus provides individual profiles that are accessible to all authors, institutions, and serial sources, hence enhancing its use and convenience. Furthermore, numerous scholarly articles have substantiated the assertion that Scopus offers a broader range of subject coverage compared to WoS. Additionally, Scopus is known for indexing a greater number of distinct sources that are not included in the WoS database [33]. The subsequent stage is the specification of the search string, encompassing all publi-cations that incorporate titles, abstracts, and keywords incorporating the term "Inter-net of Things" as a cohesive phrase, rather than as individual words. The two primary keywords selected possess the necessary qualities to effectively accomplish the objective: they are sufficiently focused and precise, avoiding excessive breadth. In the first experiment, 8215 documents from the search topic were found in the Scopus database. In addition, search results are limited to 2018 to 2022 (5 years). This research also limits the types of documents, namely journals and conference

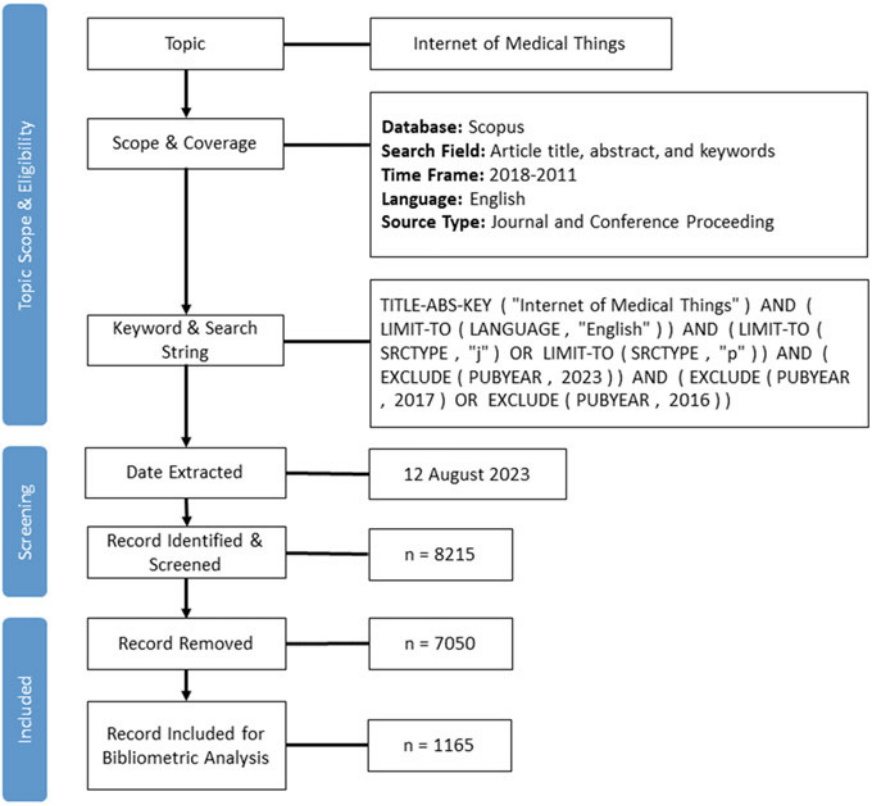


Fig. 1 The research diagram for internet of medical things: a bibliometric analysis of research publications from 2018 to 2022 [41]

proceedings. A total of 1165 articles were obtained as a consequence of the search limitation, and these articles were subsequently selected for further study. The flow chart illustrating the data collecting approach is depicted in Fig. 1.

2.2 Data Analysis

The initial step involves conducting a descriptive analysis of the data in order to assess its quality and quantity, employing established bibliometric methods. In aca-demic research, the quantity dimension is often assessed by counting the total num-ber of publications (TP). Quality dimensions, on the other hand, are assessed using a variety of measures, such as the total citations (TC) received by the publications. Furthermore, bibliometric measurements commonly incorporate the h-index as a means of forecasting future accomplishments, as opposed to conventional metrics. These

indicators are utilized at various levels, including the national level, organizational level, journal level, and individual author level. The data undergoes processing and evaluation through the utilization of the Harzing Publish or Perish (PoP) software. This process entails the extraction of Research Information System (RIS) data from the Scopus database [34–36].

The subsequent phase of the study involves the visualization of the research network in order to get insights into the interconnections of various entities, such as authors, affiliations, nations, citations, and keywords. This study examines the frequency, edge, and cluster of terms. The amalgamation of nodes, which represent keywords, and edges, which denote the interactions between these keywords, gives rise to clusters encompassing diverse research areas [37]. Nodes of larger size represent a greater frequency of occurrence in keyword visualization, whereas edges of greater thickness suggest a stronger level of linkage [38, 39]. The cluster analysis conducted in this study serves the purpose of grouping similar terms together, while distinguishing them from keywords in other groups. Its aim is to find research interests and combinations of keywords within each group. Cluster mapping is conducted using VOSviewer, a freely available software tool designed for the creation and visualiza-tion of bibliometric maps [36, 40].

3 Result and Discussion

3.1 Literature Retrieved Description

This study is limited to Internet of Medical Things research topics published between 2018 to 2022 (5 years). This research also limits the types of documents, namely journals and conference proceedings. In the end, this study obtained a total of 1165 articles that were reviewed in English. The full citation metrics for those articles are shown in Table 1.

Table 1 Metrics from this research topic

Metrics	Data
Papers	1165
Number of citations	13,411
Years	2018–2022
h-index	67

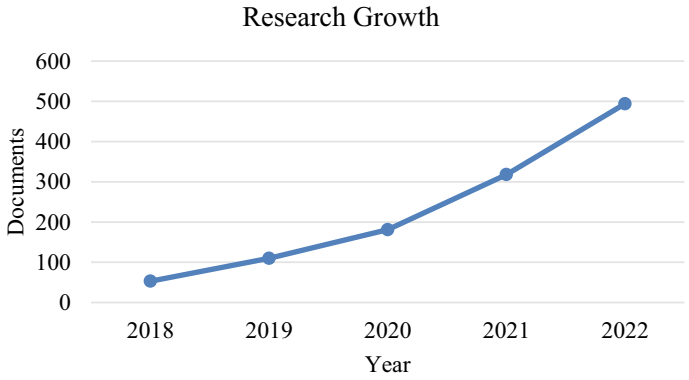


Fig. 2 Increased research growth in the last five years

3.2 Research Growth

The productivity of Internet of Medical Things research topic documents has increased throughout the year, especially research in the last five years (2018–2022). In 2018 there were 53 publications related to the topic of the Internet of Medical Things. Then in 209, the number of publications continued to increase to 110 documents. This increase in the number of studies and documents continues to increase as the years increase, 2020 (n = 181) and 2021 (n = 318). In contrast, the year 2022 exhibited the highest level of productivity, as evidenced by the publication of 503 publications. Figure 2 displays the trends in publication rates and the average number of citations per article.

3.3 Top Research Countries, Institutions, and Authors

India, China, and the United States were the most productive countries with 311, 223, and 178 articles, respectively, out of the top ten countries (Table 2). India had a doc-ument with the highest citations in 2020 with 246 citations [42]. Meanwhile, China has a document with the highest citations in 2021, which is 153 citations [43]. In the United States, namely in 2018 with 363 highest citations [44].

The variables used in this study are as follows: TP, which represents the total number of publications; NCP, which represents the number of cited publications; TC, which represents the total number of citations; and h, which represents the h-index.

According to Table 3, King Saud University has demonstrated remarkable productivi-ty by publishing a total of 47 publications. Additionally, it has achieved the greatest h-index among the institutions under consideration. The number of these

Table 2 The top ten countries contribute to the publications

Country	TP	TC	h
India	311	3876	44
China	223	2923	37
United States	178	2764	35
Saudi Arabia	162	2556	37
Pakistan	99	1538	28
United Kingdom	74	1016	19
Italy	68	1030	20
South Korea	68	711	20
Australia	51	1067	19
Malaysia	51	534	18

Table 3 Top ten affiliation contribute to the publications research

Institution	Country	TP	TC	h
King Saud University	Saudi Arabia	47	972	20
Chinese Academy of Sciences	China	29	662	15
University of North Texas	United States	27	435	13
Vellore Institute of Technology	India	25	798	15
Sejong University	South Korea	21	174	9
Instituto de Telecomunicações	Portugal	20	330	11
University of Electronic Science and Technology of China	China	19	222	7
Taif University	Saudi Arabia	19	187	11
Prince Sattam Bin Abdulaziz University	Saudi Arabia	18	137	8
Qatar University	Qatar	17	282	9

publications is higher than the second most prolific, the Chinese Academy of Sciences (n = 29). Top ten list of most prolific affiliates in Saudi Arabia.

The variables used in this study are as follows: TP, which represents the total number of publications; NCP, which represents the number of cited publications; TC, which represents the total number of citations; and h, which represents the h-index. Figure 3 shows 10 productive authors based on the number of documents published. The most documents are Moshanty, S.P., and the second is Kougianos, E.

3.4 The Most Active Journals and the Most Cited Articles

The number of articles published in various journal sources has experienced ups and downs in the number of documents. It can be noted that each publisher has a different number of documents each year from 2018 to 2022 (Fig. 4). The most productive

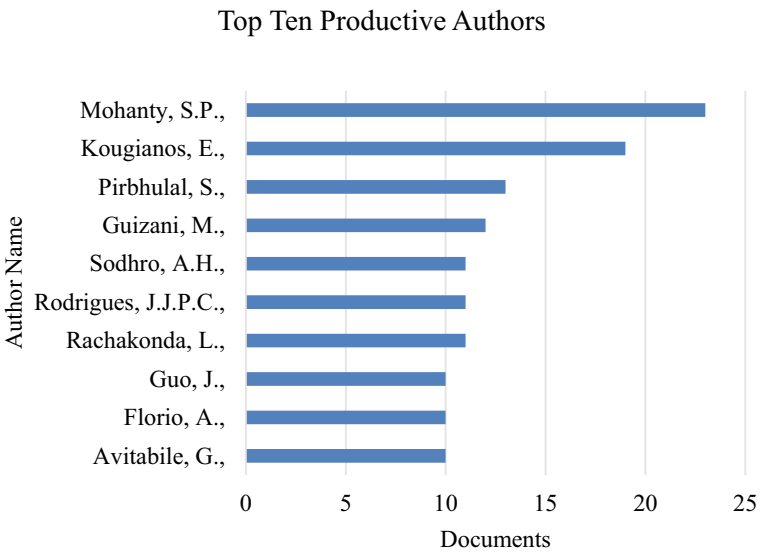


Fig. 3 The top ten productive authors

journal publisher is IEEE Access with 84 documents from 2018 to 2022. The second most is from the IEEE Internet Of Things Journal with 59 documents.

Researchers also mapped the five articles with the highest citations based on the search topic "Internet of Medical Things" (Table 4). The highest article with the highest number of citations was published in 2018, which was 363 citations. The

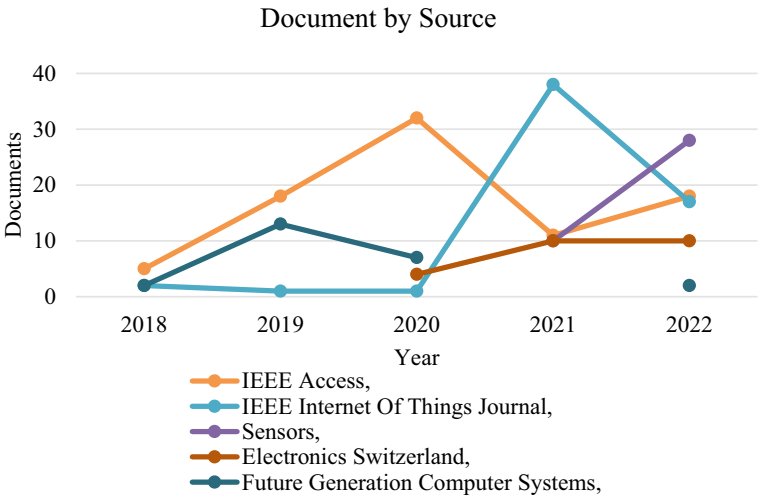


Fig. 4 Document per year by source

Table 4 The recommended placement for table captions is above the tables

No	Authors	Title	Citation
1	Banerjee et al.	"A blockchain future for internet of things security: a position paper" [44]	363
2	Nguyen et al.	"Blockchain for Secure EHRs Sharing of Mobile Cloud Based E-Health Systems" [45]	280
3	Khezr et al.	"Blockchain technology in healthcare: A comprehensive review and directions for future research" [46]	249
4	Swarna Priya et al.	"An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture" [42]	246
5	Gatouillat et al.	"Internet of Medical Things: A Review of Recent Contributions Dealing with Cyber-Physical Systems in Medicine" [47]	233

article with the second highest citation in 2019 was 280 citations, as well as the third highest in the same year.

3.5 Keyword Network Analysis of Internet of Medical Things Publications

This study systematically identified keywords based on the topic of Internet of Medical Things using Vosviewer (Fig. 5). The researcher chose Title and abstract fields to extract using the BinaryCounting calculation method. Of the 24,832 terms, 665 meet the threshold and the default choice is to select the 60% most relevant terms.

The results of network visualization found that there were 4 clusters based on their respective colors. The first cluster (135 items) with a red Covid theme. The second cluster (112 items) with the Security theme has a green color. The third cluster (105 items) with the Accuracy theme has a blue color. The fourth cluster (47 items) with the simulation result theme has a yellow color. The larger the color nodes indicate that the keyword is most often mentioned in the metadata that has been collected [48, 49]. Meanwhile, the smaller the node, it means that the keyword is mentioned less often in the metadata paper that has been collected.

3.6 Author Network Analysis of Internet of Medical Things Publications

Researchers also tried to map the author network based on the metadata obtained. Is there a relationship between authors of the research that has been done? There are 3574 authors, and 110 meet the threshold. The results showed that there were 11 clusters with different colors that were interconnected between author names (Fig. 6).

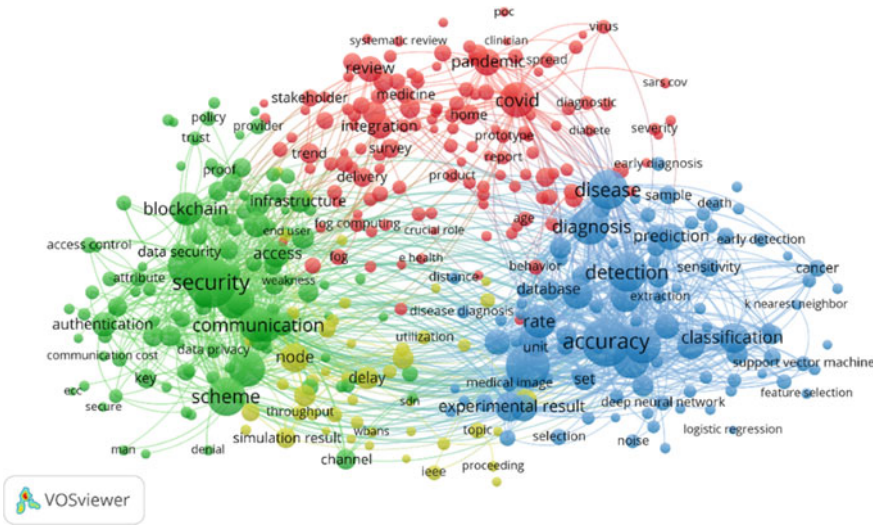


Fig. 5 Co-keyword network in Internet of Medical Things in 2018–2022

The network can show several things, such as from authors who have done research together or they come from the same affiliation [50].

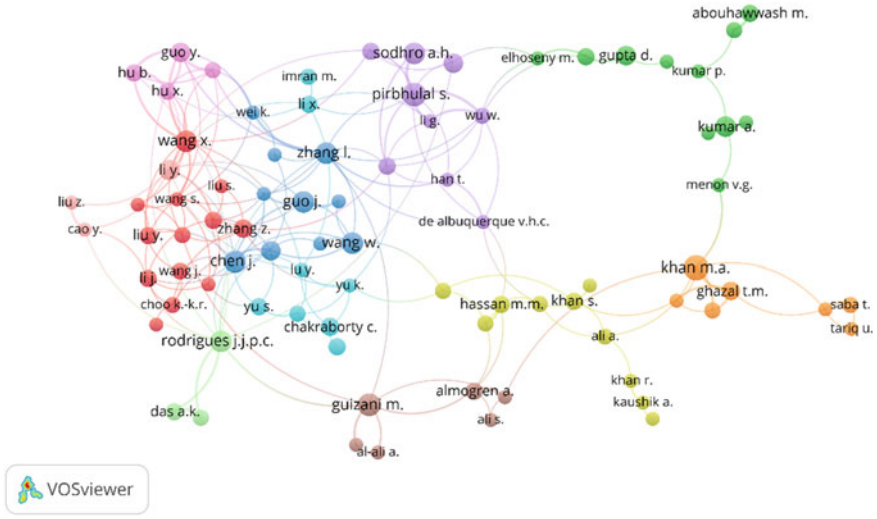


Fig. 6 Author network analysis with Vos viewer

4 Conclusion

"Internet of Medical Things: A bibliometric analysis of research publications from 2018 to 2022" gives a bibliometric study of research papers on the Internet of Medical Things (IoMT) from 2018 to 2022, using data from the Scopus database. The analysis included trends in publication output, top authors and institutions, and key research topics. The results show a steady increase in publication output in IoMT over the past five years, with the highest number of publications in 2022. The top authors in this field are Mohanty, S.P., Kougianos, E., and Pirbhulal, S., while the top institutions are King Saud University, Chinese Academy of Sciences, and University of North Texas. Citation analysis revealed that wearables, telemedicine, and data analytics are among the key research topics in IoMT. This bibliometric analysis provides valuable insights into the current state of IoMT research from 2018 to 2022 and can be useful for researchers, practitioners, policymakers, and funding agencies in understanding trends and priorities in this field. The leading authors and institutions identified in this analysis can be a valuable resource to collaborate with and keep abreast of the latest developments in IoMT. Overall, the paper highlights the importance of IoMT as a rapidly growing field that integrates medical devices and systems with the internet and other digital technologies, and the need for continued research and collaboration in this area.

References

1. Kawale SR, Narmadha D, Nancy RG, Mule SB, Verma BK, Reddy PCS (2022) Internet of medical things. Int J Health Sci (Qassim)
2. Alshammari HH (2023) The internet of things healthcare monitoring system based on MQTT protocol. Alexandria Eng J
3. Al-Turjman F, Nawaz MH, Ulusar UD (2020) Intelligence in the Internet of Medical Things era: a systematic review of current and future trends. Comput Commun 150:644–660. https://doi.org/10.1016/j.comcom.2019.12.030
4. Issa BA, Thabit QQ (2022) Review in IoT for healthcare in our life. Iraqi J Electr Electron Eng 18:9–20. https://doi.org/10.37917/ijeee.18.1.2
5. Pradhan B, Bhattacharyya S, Pal K (2021) IoT-based applications in healthcare devices. J Healthc Eng 2021:6632599. https://doi.org/10.1155/2021/6632599
6. Saini P, Gupta L, Malhotra S (2023) Internet of medical things: current applications, challenges and future. In: 2023 international conference on IoT, communication and automation technology (ICICAT), pp 1–5. https://doi.org/10.1109/ICICAT57735.2023.10263636
7. Forrest S, Baker K, Ketel M (2021) Internet of medical things: enabling key technologies. In: Conference proceedings—IEEE SOUTHEASTCON. Institute of Electrical and Electronics Engineers Inc., Applied Information Technology, University of Baltimore, Baltimore, MD, United States. https://doi.org/10.1109/SoutheastCon45413.2021.9401862
8. Alshehri F, Muhammad G (2021) A comprehensive survey of the Internet of Things (IoT) and AI-based smart healthcare. IEEE Access 9:3660–3678. https://doi.org/10.1109/ACCESS.2020.3047960
9. Greco L, Percannella G, Ritrovato P, Tortorella F, Vento M (2020) Trends in IoT based solutions for health care: moving AI to the edge. Pattern Recognit Lett 135:346–353. https://doi.org/10.1016/j.patrec.2020.05.016

10. Vyas S, Bhargava D (2021) Devices, systems and infrastructures for smart health BT—smart health systems: emerging trends. https://doi.org/10.1007/978-981-16-4201-2_6
11. Kakhi K, Alizadehsani R, Khosravi A, Nahavandi S, Islam SMS (2021) Challenges of internet of medical things for electronic healthcare. In: 2021 IEEE 4th international conference and workshop Óbuda on electrical and power engineering (CANDO-EPE), pp 19–24 (2021). https://doi.org/10.1109/CANDO-EPE54223.2021.9667929
12. Kumar M, Verma S, Kumar A, Ijaz MF, Rawat DB (2022) ANAF-IoMT: a novel architectural framework for IoMT-enabled smart healthcare system by enhancing security based on RECC-VC. IEEE Trans Ind Informatics 18:8936–8943. https://doi.org/10.1109/TII.2022.3181614
13. Khan MA, Din IU, Majali T, Kim B-S (2022) A survey of authentication in internet of things-enabled healthcare systems. Sensors 22. https://doi.org/10.3390/s22239089
14. Aljabr AA, Kumar K (2022) Design and implementation of Internet of Medical Things (IoMT) using artificial intelligent for mobile-healthcare. Meas Sensors 24. https://doi.org/10.1016/j.measen.2022.100499
15. Yuldashev Z, Sergeev A, Nastueva N (2021) IoMT technology as the basis of wearable online monitors for space distributed monitoring systems for pregnant women. In: 2021 wave electronics and its application in information and telecommunication systems (WECONF), pp 1–4. https://doi.org/10.1109/WECONF51603.2021.9470556
16. Sun Y, Lo FP-W, Lo B (2019) Security and privacy for the internet of medical things enabled healthcare systems: a survey. IEEE Access 7:183339–183355. https://doi.org/10.1109/ACCESS.2019.2960617
17. Piya K, Au QA, Shrestha S, Singh A, Mohd TK (2021) IoT in health care industry: a promising prospect. In: 2021 IEEE 12th annual ubiquitous computing, electronics & mobile communication conference (UEMCON), pp 466–474. https://doi.org/10.1109/UEMCON53757.2021.9666731
18. Yaacoub J-PA, Noura M, Noura HN, Salman O, Yaacoub E, Couturier R, Chehab A (2020) Securing internet of medical things systems: limitations, issues and recommendations. Futur Gener Comput Syst 105:581–606
19. Kashyap V, Kumar A, Kumar A, Hu Y-C (2022) A systematic survey on fog and IoT driven healthcare: open challenges and research issues. Electronics 11. https://doi.org/10.3390/electronics11172668
20. Hameed SS, Hassan WH, Latiff LA, Ghabban F (2021) A systematic review of security and privacy issues in the internet of medical things; the role of machine learning approaches. PeerJ Comput Sci 7:1–44. https://doi.org/10.7717/peerj-cs.414
21. Ali A, Nisar S, Khan MA, Mohsan SAH, Noor F, Mostafa H, Marey M (2022) A privacy-preserved internet-of-medical-things scheme for eradication and control of dengue using UAV. Micromachines 13
22. Shakeel T, Habib S, Boulila W, Koubaa A, Javed AR, Rizwan M, Gadekallu TR, Sufiyan M (2023) A survey on COVID-19 impact in the healthcare domain: worldwide market implementation, applications, security and privacy issues, challenges and future prospects. Complex Intell Syst 9:1027–1058. https://doi.org/10.1007/s40747-022-00767-w
23. Pise AA, Almuzaini KK, Ahanger TA, Farouk A, Pant K, Pareek PK, Nuagah SJ (2022) Enabling artificial intelligence of things (AIoT) healthcare architectures and listing security issues. Comput Intell Neurosci 2022:8421434. https://doi.org/10.1155/2022/8421434
24. Anugerah AR, Muttaqin PS, Trinarningsih W (2022) Social network analysis in business and management research: a bibliometric analysis of the research trend and performance from 2001 to 2020. Heliyon 8:e09270. https://doi.org/10.1016/j.heliyon.2022.e09270
25. Dai W, Zhang Z, Wang S, Liu C, Yao R, Shen L (2019) Exploring the knowledge structure of research on internet of medical things: bibliometric and visualization analysis (Preprint)
26. Gutiérrez-Salcedo M, Martínez MÁ, Moral-Munoz JA, Herrera-Viedma E, Cobo MJ (2018) Some bibliometric procedures for analyzing and evaluating research fields. Appl Intell 48:1275–1287. https://doi.org/10.1007/s10489-017-1105-y
27. Dwivedi R, Mehrotra D, Chandra S (2022) Potential of internet of medical things (IoMT) applications in building a smart healthcare system: a systematic review. J Oral Biol Craniofacial Res 12:302–318. https://doi.org/10.1016/j.jobcr.2021.11.010

28. Srivastava J, Routray S, Ahmad S, Waris MM (2022) Internet of medical things (IoMT)-based smart healthcare system: trends and progress. Comput Intell Neurosci 2022:7218113. https://doi.org/10.1155/2022/7218113
29. Tripathy S, Mohapatra UM, Mazumdar N (2022) IoT for smart healthcare: opportunities, challenges and technology. In: 2022 international conference on machine learning, computer systems and security (MLCSS), pp 171–175. https://doi.org/10.1109/MLCSS57186.2022.00039
30. Ahmad MO, Siddiqui ST (2022) The internet of things for healthcare: benefits, applications, challenges, use cases and future directions. In: Advances in data and information sciences. Presented at the (2022)
31. Almotairi KH (2023) Application of internet of things in healthcare domain. J Umm Al-Qura Univ Eng Archit 14:1–12. https://doi.org/10.1007/s43995-022-00008-8
32. Carmona-Serrano N, López-Belmonte J, Cuesta-Gómez J-L, Moreno-Guerrero A-J (2020) Documentary analysis of the scientific literature on autism and technology in web of science. https://doi.org/10.3390/brainsci10120985
33. Pranckutė R (2021) Web of Science (WoS) and scopus: the titans of bibliographic information in today's academic world. https://doi.org/10.3390/publications9010012
34. Maulana FI, Atmaji LT, Azis B, Hidayati A, Ramdania DR (2021) Mapping the educational applications of augmented reality research using a bibliometric analysis. In: 2021 7th international conference on electrical, electronics and information engineering (ICEEIE). IEEE, pp 1–6. https://doi.org/10.1109/ICEEIE52663.2021.9616934
35. Purnomo A, Susanti T, Anisah HU, Sari AK, Maulana FI (2021) Value of M-commerce research: a bibliometric perspective. In: 2021 international conference on information management and technology (ICIMTech). IEEE, Bina Nusantara University, BINUS Business School Undergraduate Program, Entrepreneurship Department, Jakarta, 11480, Indonesia, pp 813–818. https://doi.org/10.1109/ICIMTech53080.2021.9534928
36. Maulana FI, Adi PDP, Lestari D, Ariyanto SV, Purnomo A (2023) The scientific progress and prospects of artificial intelligence for cancer detection: a bibliometric analysis. In: 2023 international seminar on intelligent technology and its applications (ISITIA), pp 638–642. https://doi.org/10.1109/ISITIA59021.2023.10221162
37. Dhamija P, Bag S (2020) Role of artificial intelligence in operations environment: a review and bibliometric analysis. TQM J. 32:869–896. https://doi.org/10.1108/TQM-10-2019-0243
38. Purnomo A, Firdaus M, Prasetyo YT, Rosyidah E, Persada SF, Nathaniela TC, Maulana FI (2024) E-commerce on startup: a systematic literature review. In: Computer science and engineering in health services. Presented at the (2024)
39. Sanjaya AN, Purnomo A, Prasetyo YT, Tho C, Maulana FI, Persada SF (2023) Marketing in wireless communication: a systematic review. In: Zhao J (ed) Wireless and satellite systems. Springer Nature Switzerland, Cham, pp 85–93
40. van Eck NJ, Waltman L (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 84:523–538. https://doi.org/10.1007/s11192-009-0146-3
41. Zakaria R, Ahmi A, Ahmad AH, Othman Z (2021) Worldwide melatonin research: a bibliometric analysis of the published literature between 2015 and 2019. Chronobiol Int 38:27–37. https://doi.org/10.1080/07420528.2020.1838534
42. Swarna Priya RM, Maddikunta PKR, Parimala M, Koppu S, Gadekallu TR, Chowdhary CL, Alazab M (2020) An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture. Comput Commun 160:139–149. https://doi.org/10.1016/j.comcom.2020.05.048
43. Ning Z, Dong P, Wang X, Hu X, Guo L, Hu B, Guo Y, Qiu T, Kwok RYK (2021) Mobile edge computing enabled 5G health monitoring for internet of medical things: a decentralized game theoretic approach. IEEE J Sel Areas Commun 39:463–478. https://doi.org/10.1109/JSAC.2020.3020645
44. Banerjee M, Lee J, Choo K-KR (2018) A blockchain future for internet of things security: a position paper. Digit Commun Netw 4:149–160. https://doi.org/10.1016/j.dcan.2017.10.006

45. Nguyen DC, Pathirana PN, Ding M, Seneviratne A (2019) Blockchain for secure EHRs sharing of mobile cloud based E-health systems. IEEE Access 7:66792–66806. https://doi.org/10.1109/ACCESS.2019.2917555
46. Khezr S, Moniruzzaman M, Yassine A, Benlamri R (2019) Blockchain technology in healthcare: a comprehensive review and directions for future research. Appl Sci 9. https://doi.org/10.3390/app9091736
47. Gatouillat A, Badr Y, Massot B, Sejdic E (2018) Internet of medical things: a review of recent contributions dealing with cyber-physical systems in medicine. IEEE Internet Things J 5:3810–3822. https://doi.org/10.1109/JIOT.2018.2849014
48. Maulana FI, Zaın MY, Lestari D, Purnomo A, Adi PDP (2023) Mapping the literature of artificial intelligence in medical education: a scientometric analysis. In: Inventive communication and computational technologies. Presented at the (2023)
49. Mboi N, Syailendrawati R, Ostroff SM, Elyazar IR, Glenn SD, Rachmawati T, Nugraheni WP, Ali PB, Trisnantoro L, Adnani QES, Agustiya RI, Laksono AD, Aji B, Amalia L, Ansariadi A, Antriyandarti E, Ardani I, Ariningrum R, Aryastami NK, Djunaedi D, Efendi F, Fauk NK, Ghozali G, Handayani N, Harapan H, Hargono A, Harso AD, Ikawati HD, Indriasih E, Karlina K, Kosen S, Kusnali A, Kusuma D, Lumbantoruan M, Lusiana M, Martini S, Meilinda M, Mubasyiroh R, Musadad DA, Ningrum DNA, Nuraini S, Nurlaela N, Pangaribuan HU, Pawitaningtyas I, Purnomo A, Puspasari HW, Puspasari N, Rahmawaty S, Rangkuti H, Roosihermiatie B, Rosita T, Rustika R, Samodra YL, Samosir JV, Siswanto S, Smith KH, Sudaryanto A, Sugianto S, Suparmi S, Susilarini NK, Tarigan IU, Ticoalu JHV, Trihandini I, Wahidin M, Warouw TS, Widyastuti R, Wulandari RD, Hay SI, Murray CJL, Mokdad AH (2022) The state of health in Indonesia's provinces, 1990–2019: a systematic analysis for the global burden of disease study 2019. Lancet Glob Heal 10:e1632–e1645. https://doi.org/10.1016/S2214-109X(22)00371-0
50. Lestari D, Maulana FI, Persada SF, Adi PDP (2023) Machine learning for perinatal complication prediction: a systematic review. Inventive communication and computational technologies. Presented at the (2023)

Integration of a Reverse Vending Machine Sensing System in Sorting and Detecting Plastic Bottle Waste



Juansah, Mohamad Ramdhani, and Dien Rahmawati

Abstract A reverse vending machine is a device that uses garbage, such as used plastic bottles, to produce rewards when it is fed into the system one unit at a time. The loadcell sensor module HX711 and light-dependent resistor (LDR) sensor, as well as the ultrasonic sensor HCSR04, are used in the system's design. The system also uses the Arduino Mega 2560 microcontroller. This study aims to design an embedded system to assist in the identification of plastic bottle waste based on volume, transparency, and weight, to observe the test findings, and to contrast the test results with earlier research on the plastic bottleneck detection system. The study approach involves integrating the many parts of the plastic bottle garbage detecting systems into a single system by building system integration. The Plastic Bottle Garbage Detection System developed and installed in the Reverse Vending Machine (RVM) provided detection results of up to 95.33% of the testing results of 33 sample plastic bottler garbs in various sizes tested ranging from 160 to 1500 ml with a different number of fruits per sample. On each sample tested, testing was done up to 30 times to see how many detections there were. As a reference and benchmark for the development of plastic bottle waste detection systems using straightforward components with better results in the future, this improved detection result is anticipated to be a fundamental focus for the field of sensing systems and electronics education.

Keywords Reverse Vending Machine · Light Dependent Resistor · Load Cell · Ultrasonic · Plastic Bottle Waste

1 Introduction

A reverse vending machine is one example of an innovation in technology that sorts waste automatically, the framework for the implementation of a standard recycle bin to update the waste management system, innovate, and be creative [1]. Reselling

Juansah · M. Ramdhani (✉) · D. Rahmawati
School of Electrical Engineering, Telkom University, Bandung, Indonesia
e-mail: mohamadramdhani@telkomuniversity.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_28

plastic bottle waste to collectors or waste banks gives it marketable or financial value. A brief discussion with a scavenger revealed that plastic bottle garbage is a valuable source of gold, because it is easily accessible and easily sold. One Arduino mega microcontroller, one loadcell sensor, one hc-sr04 sensor, and one light dependent resistor sensor are used in the study. That cost should be decreased by this design. Reverse vending machines can be made more affordable [2, 3].

Based on previous studies related to the Reverse Vending Machine (RVM), such as the research titled "Development of Reverse Vending Machine (RVM) Framework for Implementation to a Standard Recycle Bin" conducted in 2016 by Razali Tomari and his team from the Faculty of Electrical Engineering and the Faculty of Civil Engineering at Tun Hussein University, Malaysia in 2017, which focused on implementing an embedded system in the Reverse Vending Machine (RVM) integrated into conventional automated waste bins based on Arduino, detecting input from plastic bottles and paper. The study achieved a tracking success rate of 92.28% with an average sensing time of 36 s [1]. The researchers aim to focus on creating a sensing system specifically designed to detect and sort plastic bottle waste.

Then, in 2018, Aditya Gaur and his team from Sharda University, Knowledge Park, India, published a research paper titled "A Simple Approach to Design a Reverse Vending Machine," which focused on the creation of a Reverse Vending Machine (RVM) and provided insights into the algorithms and methods used in building such a machine [4]. The study offered limited information about the system used, particularly the detection and sorting system it employed [5].

In 2021, a research study presented by Noor Hasyimah Abu Rahim from the University of Kuala Lumpur titled "Development of PET bottle shredder reverse vending machine" emerged, which focuses on the creation of a Reverse Vending Machine (RVM) with a primary emphasis on designing a plastic bottle waste shredder [6].

Based on the aforementioned studies, the research has primarily focused on variations [7] or models of the Reverse Vending Machine (RVM) created, where the detection system used was not discussed in detail. According to the researchers, it is crucial to understand the specific detection systems employed, given that the created Reverse Vending Machines (RVM) all converge toward a single objective [8], which is detecting plastic bottle waste. In light of this, through this journal, the researchers aim to address the subject of plastic bottle waste detection systems within Reverse Vending Machines (RVMs) using simple electronic components. It is hoped that this research can serve as a reference and benchmark for the development of plastic bottle waste detection systems using basic electronic components, contributing to the field of embedded systems and electronics education.

2 Materials and Methods

The goal was to develop a machine that could automatically classify six distinct categories of plastic bottles for reverse vending machine [3, 9]: 1. small transparent plastic bottle, 2. small colored plastic bottle, 3. medium transparent plastic bottle, 4. medium colored plastic bottle, 5. large transparent plastic bottle, 6. large colored plastic bottle. In the section that follows, theories and ideas relevant to development are also discussed.

2.1 Light Dependent Resistor (LDR)

It is a sensor that can read the value of Lx (light intensity) and receive it. This sensor will read the Lx value and has strong light sensitivity [10]. By using an analog reading of the light intensity value, which is blocked by plastic bottle waste, light light dependent resistor in a plastic bottle can be detected [10]. When the light dependent resistor receives a higher Lx value, it indicates that plastic bottle are being recognized as colored or thick, whereas when the Light Dependent Resistor (LDR) receives a lower Lx value, it indicates that plastic bottles are being detected as transparent [11].

2.2 Loadcell

Sensors with value calculations include a strong analog readout. For the purpose of converting the reading into grams [12]. The weight measurement is one of the factors used to determine how many coins will be provided to the user and is intended to stop plastic bottle waste that is still partially filled with water from entering the system [1, 2].

2.3 HC-SR04

The system has an ultrasonic sensor that measures the distance between it and the closest object being crossed in order to detect plastic bottle waste. While functioning at frequencies exceeding human hearing, ultrasonic sensors are fundamentally sound [13, 14]. This sensor's operation is based on the idea of sound wave reflection, allowing it to be used to determine the position (distance) of an item with a specific frequency [15]. Sound wave can be converted by ultrasonic sensors into many units, including height, speed, and distance. This sensor's detected range, which is approximately 2 cm to a range of 400–500 cm with a resolution of 1 cm, is an advantage. An inexpensive ultrasonic is the HCSR04 sensor [15].

2.4 Microcontroller

A microcomputer chip is a physically represents an integrated circuit (IC). It is cheap and doesn't require highly complicated computations like a PC, making it a common choice for tiny systems [13, 16]. Even though it is physically small, a microcontroller plays a crucial role in a device's functionality. Typically, a microcontroller is found in devices like remote controls, DVD players, keyboards, etc. [17–19]. There is a port on the microcontroller that can be used for read data. The purpose of the input port is to read the data that will be added to the microcontroller [20]. As the main controller or command center of the circuit, the Arduino microcontroller performs this role. The Arduino microcontroller processes every input it receives and then determines the output that has been programmed there [1].

2.5 System Block Diagram

In Fig. 1 the Arduino Mega 2560 is powered by 12 V and a 5-amp current in the design method presented. Three distinct types of sensors are powered by and controlled by the Arduino to measure various aspects of plastic bottles. A plastic bottle's degree of transparency is initially determined using an LDR (light-dependent resistor) sensor. This is accomplished by aiming a light source in the opposite direction from where an LDR sensor is located. Second, by employing ultrasonic signals, the HCSR-04 sensor is used to gauge the volume of plastic bottles. These signals are transmitted to the bottle, reflected back, and the change in signal time is utilized to calculate the volume.

Finally, its weight is directly measured using the heavy sensors. The pressure on the load cell sensor is measured for this measurement. Using this method, this design system enables precise measurements of the transparency, volume, and weight of plastic bottles, which can be utilized for a number of scientific or technical applications like quality monitoring or process control.

2.6 Flowchart

In Fig. 2 start with a standby system, there are plastic bottles of garbage that are put into the reverse vending machine. When detecting entering objects, the system will perform detection. In this process, there are three detection systems that work simultaneously, the system performs measurements of weight, color or transparency, and size detection.

In this procedure, the system will reject the removal of non-compliant plastic bottle items or waste from the reverse vending machine, with the detection findings indicating "non-plastic bottle" waste and the overweight measurement results on

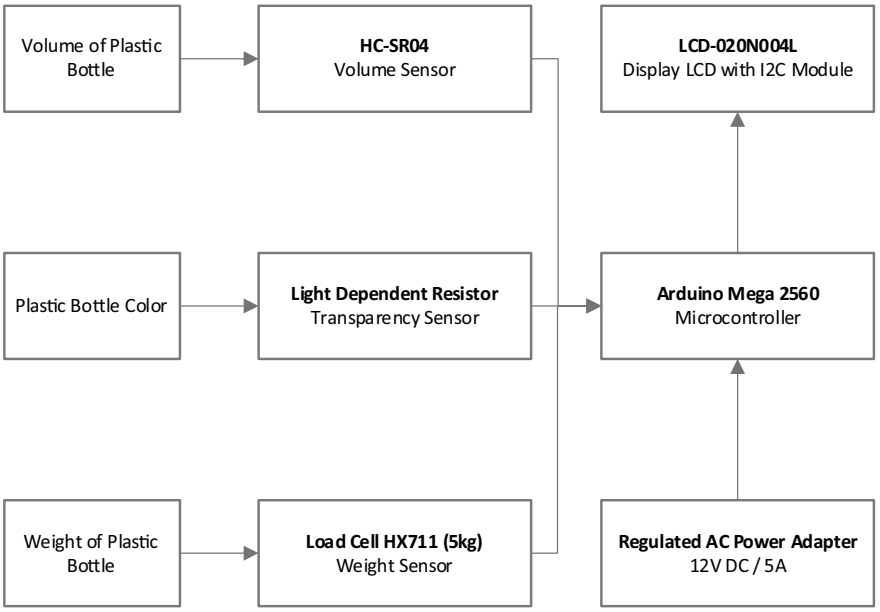


Fig. 1 Block diagram of reverse vending machine (RVM)

empty plastic bottlenecks dumped into the procedure. The machine will then go back into standby mode while it awaits someone to put the plastic bottle waste in its proper place.

3 Result

3.1 Transparency Detection of the Plastic Bottle Waste

The light sensor concept's use of a light-dependent resistor (LDR) sensor entails mounting a light-emitting diode (LED) light across the LDR sensors as in Fig. 3. The LDR sensor then reads the light readings that the LED light emits and converts them into a measurement of the light intensity (Lx), which will vary depending on the type of obstacle encountered. LDR sensor usage demonstrates 4-s detection times [10, 11].

The testing was done using a variety of conventional plastic bottle brands that are available in Indonesia. It was done using a single sample of plastic bottles of various brands and sizes, as well as other items besides plastic garbage bottles, by comparing the results of sensing and validating them through actual observation as in Table 2. The following are the outcomes of the sensing (Table 1).

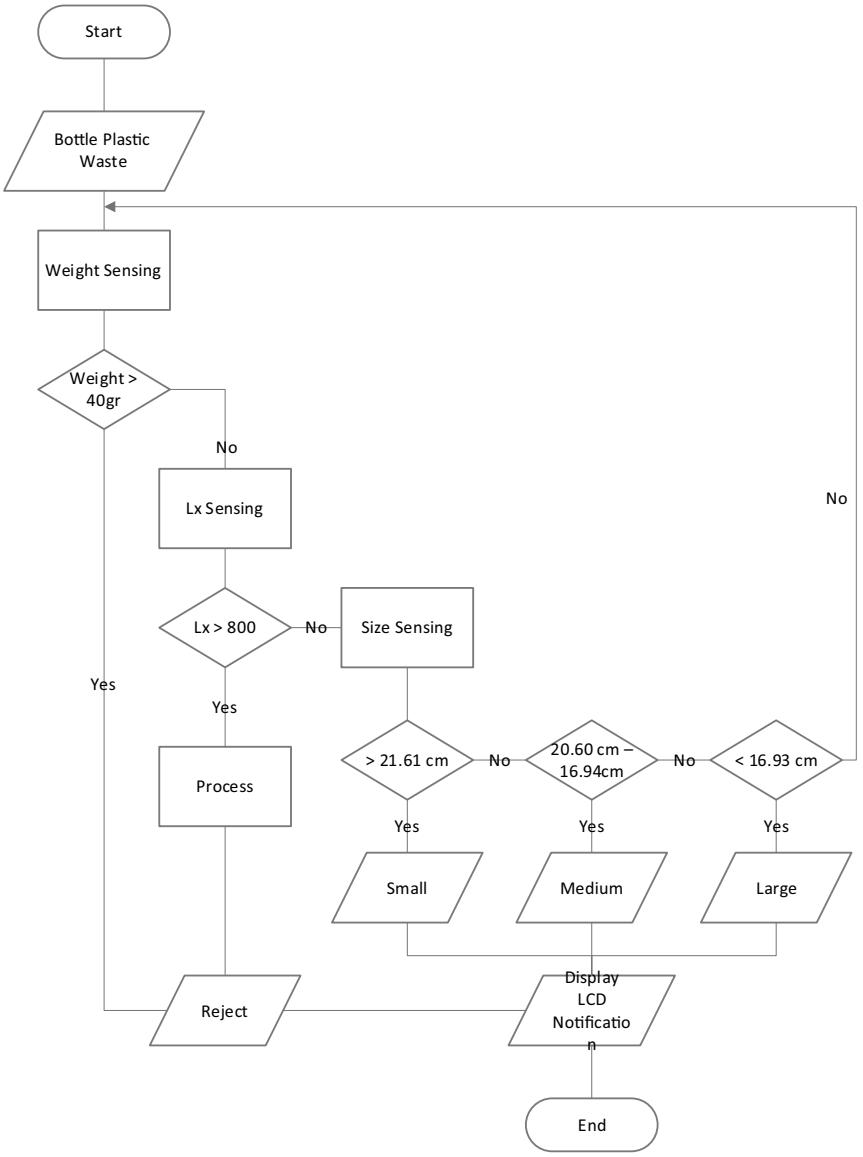


Fig. 2 Flowchart of the reverse vending machine detection system

The values of the light intensity (Lx) range from the lowest to the highest in the study of the LDR Detection (light sensor concept) outcomes. Higher values of the light intensity reported in the result indicate that the object is not transparent since the light can't pass through it.

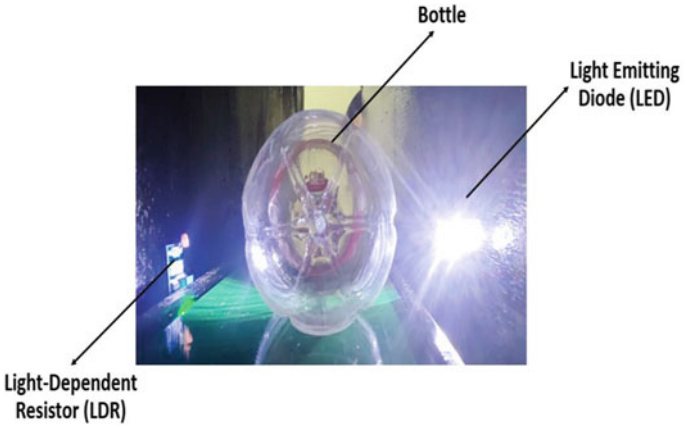


Fig. 3 Detection transparent of the plastic bottle

Table 1 Materials and light intensity (Lx)

Material type detected	Light intensity value (Lx)	Observation verification
No material detected	0	There is no test
Transparent	<149	Transparent
Color transparent	150–799	Colored
Not transparent	>800	Concentrated

Utilizing LDR (light detection) for data processing the sensor concept demonstrates a 100% correlation between the sensor detection and the observational comparison of the actual tested sample. A relevant limit between each type of sample to be evaluated is the light intensity value (Lx).

3.2 Weight Detection of the Plastic Bottle Waste

The testing was done on a variety of conventional plastic bottle brands available on the Indonesian market. For each test, a sample of plastic bottles from various brands and sizes was used, and the loadcell sensor was repeatedly measured. Three separate tests and scale readings were made, and the results frequently showed up on the loadcell sensors and digital scale as in Table 3. The measurement outcomes are as follows:

The accuracy level of detection provided by mass detection (load cell module Hx711) in carrying out the weight measurement of objects or plastic bottle waste entering the reverse vending machine system will be examined in the analysis of the first detection findings [21]. As can be seen from the test results, which show weight

Table 2 Data comparison of sensing values light intensity (Lx) and bottle color

No	Brand	Volume (ml)	Sensing result (Lx)	Material type
1	ABC Kopi Susu	200	80.50	Transparent
2	Alang Sari	300	87.00	Transparent
3	Aqua	330	80.50	Transparent
4	Aqua	600	75.00	Transparent
5	Aqua	1500	73.00	Transparent
6	Bear Brand	189	863.00	Not Transparent
7	Cimory Yogurt	240	84.50	Transparent
8	Cleo	220	93.00	Transparent
9	Cleo	750	249.00	Colored
10	Coca Cola	390	76.00	Transparent
11	Coca Cola	1000	72.50	Transparent
12	Fanta	250	79.50	Transparent
13	Fruit Tea	500	76.50	Transparent
14	Kopiko	250	76.00	Transparent
15	Lasegar	320	885.50	Not Transparent
16	Le Minerale	330	90.00	Transparent
17	Le Minerale	1500	61.00	Transparent
18	Milku	200	211,00	Colored
19	My Tea	320	79.00	Transparent
20	Nu	450	73.00	Transparent
21	Nutri Boost	300	89.00	Transparent
22	Pocari Sweat	500	71.50	Transparent
23	Pokka Greentea	450	84.00	Transparent
24	Pristine	400	93.00	Transparent
25	S-Tee	390	82.00	Transparent
26	Scarlett	100	84.00	Transparent
27	Se+	591	73.00	Transparent
28	Super O_2	600	205.00	Colored
29	Teh Botol	350	80.00	Transparent
30	Teh Pucuk	350	84.00	Transparent
31	Thai Milktea	310	90.00	Transparent
32	Vit	550	78.00	Transparent
33	Vixal	160	838.00	Not Transparent

Table 3 Comparison of sensing data and digital scales

No	Brand	Volume (ml)	Measurement result		
			Loadcell (gr)	Digital scales (gr)	Accuracy %
1	ABC Kopi Susu	200	27.95	19.27	68.94
2	Alang Sari	300	32.14	33.16	96.92
3	Aqua	330	26.50	11.05	41.70
4	Aqua	600	20.41	16.44	80.55
5	Aqua	1500	21.17	30.33	69.80
6	Bear Brand	189	34.17	31.46	92.07
7	Cimory Yogurt	240	29.77	17.57	59.02
8	Cleo	220	21.70	4.81	22.17
9	Cleo	750	27.70	31.18	87.14
10	Coca Cola	390	29.44	23.24	78.94
11	Coca Cola	1000	32.20	40.82	78.88
12	Fanta	250	22.05	12.19	55.28
13	Fruit Tea	500	27.58	33.73	81.77
14	Kopiko	250	20.33	19.27	94.79
15	Lasegar	320	27.12	13.89	51.22
16	Le Minerale	330	27.51	14.17	51.51
17	Le Minerale	1500	27.94	31.18	89.61
18	Milku	200	28.00	16.44	58.71
19	My Tea	320	24.52	19.56	79.77
20	Nu	450	30.84	22.67	73.51
21	Nutri Boost	300	26.00	24.48	93.77
22	Pocari Sweat	500	27.05	22.11	81.74
23	Pokka Greentea	450	23.35	27.49	84.94
24	Pristine	400	20.05	17.29	86.23
25	S-Tee	390	29.65	19.27	64.99
26	Scarlett	100	25.73	22.11	85.93
27	Se+	591	27.79	31.46	88.33
28	Super O_2	600	30.80	30.90	99.68
29	Teh Botol	350	30.35	25.23	83.13
30	Teh Pucuk	350	24.14	20.97	86.87
31	Thai Milktea	310	32.27	23.81	73.78
32	Vit	550	21.21	13.04	61.48
33	Vixal	160	34.51	28.91	83.77
The average value of accuracy					75.36

measurement results with an error range (%) of 0.32–77.83% and an average error result of 24.64%, the weight detection result is at 75.36% of 100% when compared to digital scales and mass detection (load cell module Hx711).

3.3 Size Measurement of the Plastic Bottle Waste

Ultrasonic waves are used to measure the separation between objects with sensors so that the size of the plastic bottle waste can be determined. The HC-SR04 sensor is used to measure the height of plastic bottle waste in a laying state, and the results of the sensing will provide a measured frequency value. The ultrasonic sensor HC-SR04 can measure an object's altitude at the system entry. In this system, it is assumed that the larger the object or bottle (around 1500 ml) the smaller the sensor's reading. This means that the smaller the size of the plastic bottle identified (around 100 ml), the larger the measurement results obtained [22, 23].

On the source code library program display, a number of defining programs have been connected to the microcontroller's digital pin. The component is placed in a higher position, face down on a pile of plastic bottle waste, and the test is conducted there. The object and the HC-SR04 ultrasonic sensor are 25.2 cm apart as in Fig. 4 [24].

By Measuring the distance between the sensor and the plastic bottle waste, the HC-SR04 ultrasonic sensor measures distance as in Table 5. To make the sensor testing process easier, the distance is calculated in centimeters (cm). This is to make it clear that the wavelength happens at a distance that is twice as high. The HC-SR04 ultrasonic sensor's test results against distance measurement in estimating the volume of plastic bottle waste through high measurements are as follows (Table 4).

When measuring the height of plastic bottle waste, the findings of the HC-SR04 sensor rangefinder were compared, and found to be quite accurate. The suitability of the result is 96.51% with the average results inaccuracy (the average calculation is 3.49%) when comparing the data validation results of sensing and measuring using a

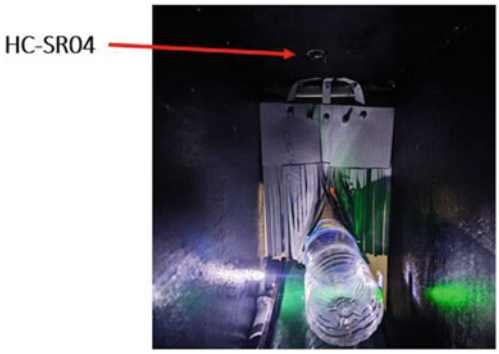


Fig. 4 The working of size measurement of the plastic bottle

Table 4 Classification of size

Distance (cm)	Size type
<19.16	Large
19.17–21.60	Medium
>21.61	Small

bar. The use of the ultrasonic sensor HC-SR04 with a presentation of sensing results of 96.51% produced an excellent result where these sensing results exceeded the reference measurement limits of plastic bottle waste [21].

3.4 Integration of All Sub Sensors

For the purpose of detecting plastic bottle waste, the entire sensor has been combined into one as in Fig. 5. To detect items that will be received by the reverse vending machine, it is already possible to detect the transparency, weight, and size of plastic bottle waste [25].

The reverse vending machine can accept an object when the entire plastic bottle waste system considers it to be in compliance with the specifications, but if it does not, the system rejects it or makes a mistake. For example, large bottles that are mistakenly identified as medium bottles or vice versa, or colorful bottles that are mistakenly identified as transparency bottles, are examples of detection errors.

A test was conducted using 33 waste plastic bottles, with each bottle unit being tested 30 times as in Table 6. The results of the calculation and presentation of success are shown in Table 6.

The entire sub's integrity test results were successful, as seen by the test results, which demonstrate an average accuracy level of 95.33% and a range of accuracy between 80 and 100%. Demonstrating that detection on the reverse vending machine is sufficient to be used and implemented.

4 Discussion

Based on the overall results of testing 33 samples of plastic bottle waste with various brands and sizes, each consisting of one unit per sample, tested 30 times to assess the success rate of the plastic bottle waste detection system on the Reverse Vending Machine (RVM), an average accuracy score of 95.33% was obtained. The detection accuracy values ranged from 80 to 100%, with a sensing time of 4 s for all tested samples of plastic bottle waste. During the testing process, some samples showed excellent detection results with a 100% accuracy rate, such as the 300 ml Alang Sari bottle, 200 ml ABC coffee bottle, 240 ml Cimory bottle, 250 ml Fanta bottle, and several other bottles with volumes ranging from 200 to 591 ml [22].

Table 5 Comparison of sensing data and rules measurements

No	Brand	Volume (ml)	Measurement result			Size type
			Load cell (gr)	Digital scales (gr)	Accuracy %	
1	ABC Kopi Susu	200	22.17	20.80	93.82	Small
2	Alang Sari	300	20.74	20.30	97.88	Medium
3	Aqua	330	21.01	20.20	96.14	Medium
4	Aqua	600	20.73	19.40	93.58	Medium
5	Aqua	1500	17.24	17.50	98.51	Large
6	Bear Brand	189	21.59	20.80	96.34	Medium
7	Cimory Yogurt	240	21.59	21.30	98.66	Medium
8	Cleo	220	21.30	21.40	99.53	Medium
9	Cleo	750	17.24	16.60	96.29	Large
10	Coca Cola	390	21.30	20.10	94.37	Medium
11	Coca Cola	1000	17.54	17.30	98.63	Large
12	Fanta	250	22.98	20.50	89.21	Small
13	Fruit Tea	500	20.90	19.60	93.78	Medium
14	Kopiko	250	20.72	20.40	98.56	Medium
15	Lasegar	320	20.93	19.50	93.17	Medium
16	Le Minerale	330	20.09	18.10	90.09	Medium
17	Le Minerale	1500	18.19	17.50	96.21	Large
18	Milku	200	21.01	20.80	99.00	Medium
19	My Tea	320	21.25	20.10	94.59	Medium
20	Nu	450	20.42	19.50	95.49	Medium
21	Nutri Boost	300	21.01	20.20	96.14	Medium
22	Pocari Sweat	500	19.14	19.30	99.17	Medium
23	Pokka Greentea	450	20.42	19.60	95.98	Medium
24	Pristine	400	21.01	20.60	98.05	Medium
25	S-Tee	390	21.01	20.30	96.62	Medium
26	Scarlett	100	22.75	22.80	99.78	Small
27	Se+	591	20.09	19.30	96.07	Medium
28	Super O_2	600	19.46	19.10	98.15	Medium
29	Teh Botol	350	20.74	19.80	95.47	Medium
30	Teh Pucuk	350	20.09	20.20	99.46	Medium

(continued)

Table 5 (continued)

No	Brand	Volume (ml)	Measurement result			Size type
			Load cell (gr)	Digital scales (gr)	Accuracy %	
31	Thai Milktea	310	20.42	20.70	98.65	Medium
32	Vit	550	20.10	19.70	98.01	Medium
33	Vixal	160	21.01	20.90	99.48	Medium
The average value of accuracy					96.51	

However, in some tested bottle samples, it was found that there were a few bottles with the lowest sensing results, with an 87% accuracy rate. Out of 30 tests, they were detected successfully 25 times (corresponding to the type of bottle inserted) and failed 5 times (the inserted bottle waste was not detected properly). This was the case with the 500 ml Pocari Sweat bottle, which should have been classified as a medium-sized transparent bottle, but the system had inconsistent readings, sometimes categorizing it as a small transparent bottle and at other times as a colored bottle. Similar inconsistencies in detection results were also observed with the Thai milk tea bottle, which had an 80% accuracy rate. In the cases of failure, the system tended to detect the plastic bottle waste as the colored type when it should have been classified as transparent.

Several factors influencing the inaccuracy of overall test results include the integration of the detection system created, which largely relies on the reading of light intensity values (lx) from the LDR sensor, where the light intensity value (lx) serves as a reference for classifying plastic bottle waste and the color of plastic bottle waste based on the detected transparency value. The integration of the plastic bottle waste sensing system utilizing the LDR sensor [10], HX711 load cell module, and ultrasonic sensor forms a unified plastic bottle waste detection system, with limited information and literature research related to plastic bottle waste sensing system research, whether implemented separately or within a Reverse Vending Machine (RVM). One of the weaknesses found in this type of sensing system integration is the low sensitivity of light intensity values in distinguishing transparent plastic bottle waste from transparent plastic bottle waste still covered by labels or brand markings, resulting in relatively low success rates compared to other types of plastic bottle waste with semi-transparent labels. Therefore, researchers can set specific light intensity value (lx) limits for certain transparent plastic bottle waste brands with labels. The researcher's observations indicate that the classification of light intensity values ranges from lx 0, indicating high transparency of plastic bottle waste or objects, to values exceeding 800, indicating dense plastic bottle waste (non-transparent) as detailed in Table 1.

Considering several previous studies [1, 4, 7], the researcher compared the research results related to the plastic bottle waste detection system that had been conducted using the data obtained by the researcher. Together, we conducted a comparison based on sensors, design methods, detection time, and the accuracy

Table 6 Data of integration of detection results

No	Brand	Volume (ml)	Sensing duration (s)	Specification		Number of detection		Accuracy %
				Size	Material	Success	Fail	
1	ABC Kopi Susu	200	4	Small	Transparent	30	–	100
2	Alang Sari	300	4	Medium	Transparent	30	–	100
3	Aqua	330	4	Medium	Transparent	27	3	90
4	Aqua	600	4	Medium	Transparent	27	3	90
5	Aqua	1500	4	Large	Transparent	26	4	87
6	Bear Brand	189	4	Medium	Not Transparent	30	–	100
7	Cimory Yogurt	240	4	Medium	Transparent	30	–	100
8	Cleo	220	4	Medium	Transparent	28	2	93
9	Cleo	750	4	Large	Colored	28	2	93
10	Coca Cola	390	4	Medium	Transparent	29	1	97
11	Coca Cola	1000	4	Large	Transparent	29	1	97
12	Fanta	250	4	Small	Transparent	30	–	100
13	Fruit Tea	500	4	Medium	Transparent	25	5	83
14	Kopiko	250	4	Medium	Transparent	29	1	97
15	Lasegar	320	4	Medium	Not Transparent	30	–	100
16	Le Minerale	330	4	Medium	Transparent	25	5	80
17	Le Minerale	1500	4	Large	Transparent	29	1	97
18	Milku	200	4	Medium	Colored	26	4	87
19	My Tea	320	4	Medium	Transparent	28	2	93
20	Nu	450	4	Medium	Transparent	26	4	87
21	Nutri Boost	300	4	Medium	Transparent	29	1	97
22	Pocari Sweat	500	4	Medium	Transparent	25	5	80
23	Pokka Greentea	450	4	Medium	Transparent	28	2	93
24	Pristine	400	4	Medium	Transparent	30	–	100
25	S-Tee	390	4	Medium	Transparent	25	5	80

(continued)

Table 6 (continued)

No	Brand	Volume (ml)	Sensing duration (s)	Specification		Number of detection		Accuracy %
				Size	Material	Success	Fail	
26	Scarlett	100	4	Small	Transparent	30	–	100
27	Se +	591	4	Medium	Transparent	30	–	100
28	Super O2	600	4	Medium	Colored	27	3	90
29	Teh Botol	350	4	Medium	Transparent	27	3	90
30	Teh Pucuk	350	4	Medium	Transparent	29	1	97
31	Thai Milktea	310	4	Medium	Transparent	24	6	80
32	Vit	550	4	Medium	Transparent	29	1	97
33	Vixal	160	4	Medium	Not Transparent	29	1	97
The average value of accuracy								95.33

of the plastic bottle waste detection system in the Reverse Vending Machine (RVM) designed by the researcher with several previous studies. The data from this comparison are summarized in Table 7.

The comparison results in Table 7 show that the research conducted on the design of an integrated plastic bottle waste detection system yielded promising outcomes. In the initial study by Tomari and the team in 2017 regarding the implementation of the Reverse Vending Machine (RVM) design as an automated trash receptacle, the integration of paper and bottle waste detection systems achieved an accuracy of 92.28% with a sensing speed of 36 s [1]. In 2018, a study was published on the

Table 7 Comparison with past studies

	Sensor	Work method	Time detection (s)	Performance
Proposed study	Ultrasonic HCSR04, Loadcell module Hx711, Light Dependent Resistor (LDR)	Integrated sensing system design	4 s	95.33%
Previous study 1	RFID, PIR Sensor, Loadcell	RFID Card Model System	36 s	92.28%
Previous study 2	–	Manufacturing simulation using Xilinx 14.5 ISE Simulator	–	–
Previous study 3	Capacitive proximity sensor	Build Framework and product design	2 s	75%

creation of a simplified Reverse Vending Machine (RVM) using Xilinx 14.5 ISE Simulator for simulation purposes. This research focused on conceptual and basic design and did not provide comprehensive information about the detection system used [4].

Moving forward to 2021, a study by Noor Hasyimah Abu Rahim from the University of Kuala Lumpur, Malaysia, regarding the Reverse Vending Machine (RVM) featuring Build Framework and product design with a Capacitive proximity sensor, resulted in various research outcomes related to the integration of the plastic bottle waste detection system in the RVM. This system achieved an accuracy rate of 75% with a detection time of 2 s [6]. In 2023, researchers presented a concept related to research on integrating a plastic bottle waste sensing system into the Reverse Vending Machine (RVM), achieving an accuracy rate of 95.33% with a detection speed of 4 s. It's worth noting that this detection speed is slightly slower (2 s) compared to the previous year's research [7]. However, it should be remembered that this represents a significant improvement when compared to the research conducted in 2017, which had a detection time of 36 s [1]. The percentage of accuracy has also increased, as previous research by Tomari and the team achieved 92.28%, Aditya Gaur and the team focused on simulations and advised considering a sensing system with good accuracy, resulting in 75% accuracy in 2021 [6]. The researchers feel that they have made significant progress in the accuracy of detection compared to previous studies, reaching 95.33%. The researchers' findings demonstrate the need to use a different sensor system model from those used in several earlier experiments in order to sense plastic bottle waste with more precision.

5 Conclusion

Based on the results and discussion, conclusions can be drawn in line with the research objectives related to the embedded system design aimed at assisting in the detection of plastic bottle waste based on volume, transparency level, and weight. This is achieved by examining the test results and comparing them with previous research on plastic bottle waste detection systems. The formation of the plastic bottle waste detection system is a physical presence consisting of three different sensing systems: the Light Dependent Resistor (LDR) sensor serves as the object detection system for plastic bottle waste by measuring the light intensity value, allowing for the determination of the object's transparency on a scale as measured in Table 1 (0–800); the HC SR04 ultrasonic sensor aids in measuring the volume of plastic bottle waste based on its size, and the load cell module HX711 sensor assists in measuring the weight of plastic bottle waste, thus serving as a reference for sorting or sensing the weight of an object entering the Reverse Vending Machine (RVM) [2, 5].

After conducting testing and measurements on each individual sensing system, the researcher proceeded to combine all components into an integrated sensing system, resulting in a plastic bottle waste detection system that can be used. After comprehensive testing with 33 objects, the final test results yielded an accuracy rate of 95.33%

with a sensing reading speed of 4 s. However, it is necessary to improve this level of precision in advance so that the sensor system applied to the reverse vending machine gives the correct result accurate approach or even 100% of the detection results.

References

1. Tomari R, Kadir AA, Zakaria WNW et al (2017) Development of reverse vending machine (RVM) framework for implementation to a standard recycle bin. In: Procedia computer science. Elsevier B.V., pp 75–80
2. Young Joung K (2021) 3D-printed load cell using nanocarbon composite strain sensor. Sensors 21:1–3
3. Balubai M et al (2017) A new approach in manufacturing of reverse vending machine. Int J Adv Eng Manage Sci (IJAEMS) 3
4. Gaur A, Priyadarshini R, Mathuria D (2018) A simple approach to design reverse vending machine precision agriculture view project development of wireless sensor network for campus monitoring view project a simple approach to design reverse vending machine
5. Watanyulersakul E (2019) The accuracy of sorting beverage cans and bottles for a reverse vending machine. ECTI Trans Comput Inform Technol 13:71–72
6. Noor Hasyimah AR, Ahmad Nor HMK (2021) Development of PET bottle shredder reverse vending machine. Int J Adv Technol Eng Explor 8:24–32
7. Rahim NHA, Khatib ANHM (2021) Development of PET bottle shredder reverse vending machine. Int J Adv Technol Eng Explor 8:24–33. https://doi.org/10.19101/IJATEE.2020.S2762167
8. Bahram Zahir Azami S, Tanabian M (2004) Automatic mobile payment on a non-connected vending machine. Can Conf ElectrComput Eng 5:731–734
9. Yoo T, Lee J-S, Kim T (2021) Dual image-based CNN ensemble model for waste classification in reverse vending machine. Appl Sci 11:1
10. Dadi V, Peravali S (2020) Optimization of light-dependent resistor sensor for the application of solar energy tracking system. SN Appl Sci 2
11. Angin DP, Siagian H, Suryanto ED et al (2018) Design and development of the trash spliter with three different sensors. J Phys Conf Ser 1007:012057. https://doi.org/10.1088/1742-6596/1007/1/012057
12. VPG Transducers (2015) Load cell technology. load cells and weight modules technical note VPGT-01. http://www.vishaypg.com/docs/11866/vpg
13. Susilo J et al (2021) Car parking distance controller using ultrasonic sensors based on Arduino uno. J Robot Control (JRC) 16:1–6
14. Qureshi MA, Aziz A, Rasool HF (2011) Design and implementation of automatic ticket system using Verilog HDL. International conference on information technology (IACSIT), pp 707–712
15. Morgan EJ HCSR04 Ultrasonic Sensor
16. Dumpayan WGP, De Mesa MLM, Yucor NDF et al (2017) Two-way powered microcontroller-based plastic bottles 'drop-and-tap' reverse vending machine with stored value system using radio frequency identification (RFID) scanner technology. In: 2017 IEEE 9th international conference on humanoid, nanotechnology, information technology, communication and control, environment and management (HNICEM). IEEE, pp 1–8
17. Amano M (2003) PET bottle system in Sweden and Japan: an integrated analysis from a life-cycle perspective. Masters' Thesis for LUMES—Lund University International Master's Programed in Environmental Science
18. Karin MF, Noor KS, Zaman HU (2016) Hardware based design and implementation of a bottle recycling machine using FPGA. IEEE
19. Halvorsen H (2018) Programming with Arduino. ISBN

20. Narendra J et al (2019) Automatic cold drink vending machine with reverse vending machine system for empty bottles and monitoring system using VLSI. Int J Innov Technol Explor Eng 8
21. Litwhiler DH (2020) Design, development, and testing of load cell accelerometers. In: Asee annual conference and exposition, conference proceedings
22. WWF Australia (2022) The lifecycle of plastics. https://www.wwf.org.au/news/blogs/the-lifecycle-of-plastics
23. Islam M, Ferdous MF, Almazee MSR et al (2019) Recyclable plastic bottles deposit and refund system (RPBDRS). In: 2019 joint 8th international conference on informatics, electronics & vision (ICIEV) and 2019 3rd international conference on imaging, vision & pattern recognition (icIVPR). IEEE, pp 45–50
24. DeGlandon K (2016) Selecting the best hydraulic pressure sensor. Drilling instruments. https://www.drillinginstruments.com/resourcesblog/2016/01/25/selecting-the-best-hydraulic-pressure-sensor/
25. (2020) The optical subsystem for the empty containers recognition and sorting in a reverse vending machine—IEEE conference publication. IEEE Conference Publication

Corrosion Resistance Improvement of Mild Steel Using *Cocos Nucifera* Leaf Extract in Seawater



W. M. Wan Syahidah, R. Rosliza, and F. Atan

Abstract Challenges related to corrosion significantly impact the durability and safety of mild steel structures when exposed to aggressive environments like seawater. This research explores the potential of *Cocos nucifera* leaf extract (CNLE) as a natural inhibitor to enhance the resistance of mild steel against corrosion in seawater. The study employs potentiodynamic polarization and linear polarization resistance techniques to assess the extract's inhibitory properties. Additionally, it conducts an analysis of inhibition efficiency and adsorption isotherms using the Langmuir model to shed light on the adsorption mechanism. The inhibition efficiency shows an increase with higher extract concentrations, eventually reaching a maximum value. The adsorption isotherm analysis validates the applicability of the Langmuir model, indicating monolayer adsorption and a strong interaction between the inhibitor and mild steel. These findings contribute to the development of environmentally friendly and cost-effective methods for enhancing the longevity and reliability of mild steel materials used in seawater applications.

Keywords *Cocos Nucifera* · Mild steel · Corrosion resistance · Potentiodynamic · Polarization · Linear polarization resistance

1 Introduction

Due its cost-effectiveness and favorable mechanical properties, mild steel is a widely employed material in marine and coastal structures [1–3]. It also plays a vital role in electrical engineering, being used in applications like power transmission towers, substation structures, and electrical enclosures. The operational lifespan of mild steel in coastal and offshore environments, where exposure to seawater is frequent and poses a substantial risk to its performance, is significantly compromised due to its high susceptibility to corrosion [4–6]. This corrosion not only jeopardizes

W. M. W. Syahidah (✉) · R. Rosliza · F. Atan
Faculty of Engineering Technology, University College TATI, Chukai, Malaysia
e-mail: syahidah@uctati.edu.my

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_29

the reliability of electrical systems but also imposes substantial maintenance and replacement costs. Hence, there is a need to explore alternative corrosion protection strategies that are both effective and environmentally friendly [7–9].

The corrosion inhibition technique encompasses the utilization of chemical substances or protective coatings specifically engineered to safeguard metals and alloys against corrosion when they are subjected to corrosive surroundings. These inhibitors function by establishing a protective shield on the metal's surface or by modifying the electrochemical mechanisms that initiate corrosion [6, 7]. Typical categories of corrosion inhibitors comprise organic compounds, inorganic compounds, and metallic coatings. Organic inhibitors can absorb onto the metal surface, while inorganic inhibitors, like phosphate and chromate compounds, form a passive film. Metallic coatings, such as zinc or cadmium plating, provide sacrificial protection by corroding in place of the underlying metal. In recent years, natural inhibitors derived from plant extracts have gained attention as potential alternatives to synthetic chemical inhibitors [5, 9–17]. The state of the art in corrosion inhibition research has revealed the effectiveness of natural plant extracts in protecting metals from corrosion in various environments.

Cocos nucifera, commonly known as the coconut palm, has been widely studied for its various applications in food, medicine, and other industries. The potential inhibitive properties of *C. nucifera* leaf extract on mild steel corrosion have been recognized in recent research.

While previous research has explored the use of plant extracts as corrosion inhibitors for metals, additional investigation is required to delve into the precise utilization of *C. nucifera* leaf extract (CNLE) for improving the corrosion resistance of mild steel in seawater [18–21]. A distinct research gap exists concerning the utilization of *C. nucifera* leaf extract as a corrosion inhibitor in seawater. This gap represents a significant opportunity to harness a waste product, the coconut leaves to simultaneously address the challenges of corrosion control and sustainable waste management. Understanding the mechanisms through which the leaf extract interacts with the mild steel surface and inhibits corrosion is crucial for practical implementation in marine environments [13].

The structure of the research article is designed to offer a thorough insight into the enhancement of mild steel's corrosion resistance when *C. nucifera* leaf extract is employed in seawater. The investigation initiates with an introductory section that establishes the context and elucidates the significance of protecting mild steel from corrosion in marine settings. It addresses the corrosion issue, underscores the importance of environmentally friendly inhibitors, and elucidates the motivation behind the utilization of CNLE.

2 Material and Method

2.1 Materials

The study utilized a mild steel sample, SS400, which measured 25 mm × 25 mm × 3 mm in size. The material's chemical composition, expressed in weight percentages, consisted of the following elements: C (0.25%), Cu (0.20%), Mn (1.03%), P (0.04%), Si (0.28%), S (0.05%), and the remainder being Fe. The coupons underwent mechanical polishing using emery papers of different grit sizes (220, 500, and 1000), facilitated by applying distilled water as a lubricant. After being polished, the samples will be cleaned in acetone and rinsed in distilled water. After that, they'll be stored over a desiccant to finish drying out in the air. Nitric acid (HNO_3) is then used to dissolve the corrosion products from the samples thoroughly. The extraction of *C. nucifera* leaf was performed using ethanol to obtain the active components of the leaf extract.

2.2 Electrochemical Measurement

Potentiodynamic polarization and linear polarization resistance are two electrochemical methodologies employed for assessing the corrosion characteristics of SS400 in diverse environments. These techniques yield valuable data concerning corrosion rate, polarization resistance, and other electrochemical parameters that aid in the evaluation of the inhibitory efficacy of substances like CNLE. The amalgamation of potentiodynamic polarization and linear polarization resistance techniques furnishes a comprehensive comprehension of the corrosion behavior and inhibitory performance of mild steel in seawater when CNLE is introduced. These electrochemical methods facilitate the quantification of the natural extract's inhibitory impact and contribute to the evaluation of its potential utility as a corrosion inhibitor in marine settings.

The working electrode (WE) specimen, obtained by cutting a square section of mild steel, possesses an exposed surface area of 3.75 cm^2 when exposed to the testing solution. A fresh working electrode is substituted for each measurement. Subsequent to immersing the working electrode (WE) in the test solution, it was submerged and permitted to stabilize at an open circuit potential. Electrochemical tests were carried out once the open circuit potential had reached a steady state. The current–potential curves in the potentiodynamic graph were measured by scanning from 100 mV to +100 mV at 1 mVs^{-1}. Appropriate electrical connections between the electrodes and the potentiostat, which is responsible for controlling the applied potential, is established. The potential scan generates a current–potential curve, also known as the Tafel plot, which provides information on the corrosion performance of mild steel.

Another electrochemical method employed for investigating the corrosion characteristics of metals is linear polarization resistance (LPR). In LPR, a slight potential disturbance is introduced to the working electrode, and the subsequent measurement records the resulting current response. LPR is based on the linear relationship between current and potential around the corrosion potential (E_{corr}). The linear LPR measurement was carried out with a scan rate of 1 mVs^{-1} from a range of -10 mV to +10 mV in relation to E_{corr}.

The inhibition efficiency (IE) of a corrosion inhibitor using the potentiodynamic polarization (PP) method can be calculated using inhibition efficiency formula [21, 25].

A higher percentage indicates a more effective corrosion inhibitor. A 100% inhibition efficiency means the inhibitor completely prevents corrosion. The adsorption nature of CNLE on SS400 was determined using the Langmuir adsorption isotherm model. The Langmuir isotherm model and the parameters of this model are calculated using Eq. 1 [12].

$$\frac{C}{\theta} = \frac{1}{K_{ads}} + C \tag{1}$$

where θ (IE/100) is the surface coverage of mild steel, K_{ads} is the adsorption equilibrium constant, and C is the concentration of CNLE. The Gibbs free energy ΔG°_{ads} calculated using the formula in Eq. 2 [22].

$$K_{ads} = \frac{1}{55.5} \exp\left(\frac{\Delta G^{\circ}_{ads}}{R.T}\right) \tag{2}$$

where R is the molar gas constant, T is the room temperature in Kevin.

3 Result

Polarization methods were conducted to estimate the impact of CNLE on corrosion reactions. The alterations observed in the polarization curve following the introduction of the inhibitor were typically employed to classify inhibitors as anodic, cathodic, or mixed based on predetermined criteria. Figure 1 represents the potentiodynamic polarization (PP) curves of SS400 in seawater in the absence and presence of CNLE. The E_{corr} and i_{corr} were calculated by fitting the anodic and cathodic Tafel lines using the "corrosion fit" analysis program of the general-purpose electrochemical system (GPES) software.

The LPR method assessed mild steel's polarization resistance (R_p) in seawater with and without the inhibitor. The linear region detected at ± 10 mV near the corrosion potential was used to calculate the R_p values. According to the data presented in Table 1, adding the inhibitor increases the R_p values of mild steel. The polarization resistance (R_p) is a measure of the mild steel's resistance to corrosion.

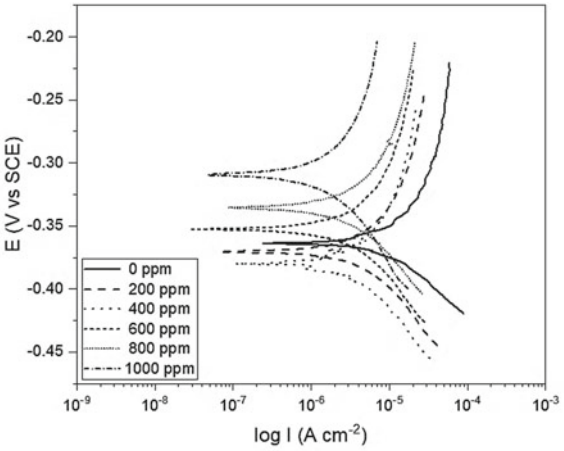


Fig. 1 PP curves at various concentrations of CNLE

Table 1 The linear polarization results at various concentrations of CNLE

Concentration (ppm)	PP			LPR
	E_{corr} (mV)	i_{corr} (μA)	C.R (10^{-2} mm/year)	R_p (kΩ)
0	−370.77	17.409	3.358	1.108
200	−384.38	8.020	1.510	3.451
400	−382.33	4.609	0.868	4.951
600	−358.96	3.540	0.666	6.220
800	−338.84	2.983	0.562	7.058
1000	−301.84	1.664	0.313	14.694

Figure 2 displays the IE results obtained from all electrochemical assessments conducted at different concentrations. The graph depicts a positive correlation between the inhibitor concentrations from 200 up to 1000 ppm and the increase in the IE of CNLE. The highest IE can exceed 90% at 1000 ppm for each method. Therefore, it can be observed that three different methods are in trend. The expansion of the inhibitor molecule coverage area on the metal surface is directly proportional to the increase in inhibitor concentration, leading to an increment in IE.

As depicted in Fig. 3, the plots revealed straight lines with an intercept of 1/K, as shown in Fig. 3. The K_{ads} and ΔG°_{ads} values obtained from the calculations are presented in Table 2 using linear fit analysis. The value of ΔG°_{ads} is correlated with the various types of adsorptions.

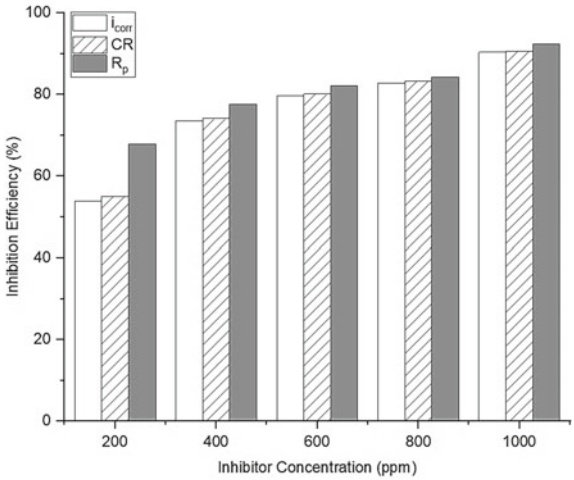


Fig. 2 Inhibition efficiency, IE (%) at various concentrations of CNLE

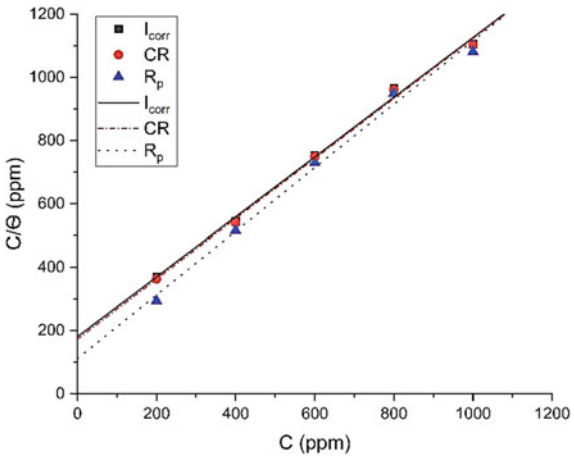


Fig. 3 The Langmuir adsorption isotherm of CNLE

Table 2 The result for linear fit analysis using Langmuir adsorption isotherm

Method		Slope, m	R^2	$K_{ads} \times 10^3$ ($dm^3\ mol^{-1}$)	ΔG_{ads} ($kJ\ mol^{-1}$)
PP	Icorr	0.946	0.996	5.540	−31.32
	CR	0.946	0.996	5.786	−31.43
LPR	Rp	1.004	0.993	8.943	−32.51

4 Discussion

The modifications observed in the cathodic and anodic Tafel slopes, as illustrated in Fig. 1, indicate that the existence of CNLE leads to a change in the anodic dissolution mechanism. Furthermore, the extract's adsorption impacts the cathodic hydrogen evolution process resulting in a significant decrease in the icorr [23]. Nitrogen and sulfur in the formulation of the inhibitor improve its overall performance, as these elements tend to form stable complexes with metal ions, thereby reducing the corrosion rate. [24]. The inhibition mechanism is thought to revolve around the adsorption of inhibitor molecules onto the metal's surface, leading to the creation of a protective film. The film acts as a shield, hindering the flow of electrons and retarding the corrosion process. The behaviour of CNLE suggests that it functions as an inhibitor of mixed type.

In Table 1, an uptick in the R_p value indicates an improvement in the metal's ability to resist corrosion when subjected to seawater containing Cocos nucifera leaf extract. A greater R_p value signifies that the inhibitor establishes a protective layer on the metal's surface, impeding the corrosion process. The creation of this protective surface layer on mild steel by CNLE is essential in impeding charge transfer and serves as an essential barrier.

The adsorption of the inhibitor onto the mild steel surface results in the formation of a thin protective coating. Consequently, in this scenario, CNLE functions as a filming corrosion inhibitor, effectively regulating the corrosion rate. Filming inhibitors work by isolating the metal from its surroundings rather than reacting with or eliminating active corrosive substances. The presence of sulfur and aromatic compounds in CNLE enhances its corrosion inhibition performance [24]. The influence of CNLE on corrosion inhibition tests for mild steel can be substantial and advantageous. The natural inhibitor found in CNLE holds promise for augmenting the corrosion resistance of mild steel when it encounters corrosive environments, such as seawater. This inhibitory characteristic of the leaf extract can be attributed to its chemical constituents, which encompass various organic compounds like phenols, flavonoids, alkaloids, and organic acids.

Corrosion inhibitors are typically adsorbed through a combination of physical and chemical mechanisms. In the realm of physisorption, it involves the electrostatic interaction that occurs between a charged corrosion inhibitor and a metal surface. On the other hand, chemical adsorption is characterized by the formation of coordination bonds between available orbitals on the metal's surface and unshared electron pairs originating from heteroatoms within the organic inhibitor [22, 25, 26]. The corrosion inhibitor molecules, a ΔG°_{ads} in Table 2 signifies that corrosion inhibitor molecules are involved in a combination of both chemical and physical adsorption interactions with the metal's surface [22, 27–29]. The linear regression coefficient (R2) closely approaches 1, signifying a strong correlation between the model and the experimental data. The trends in each method show that CNLE adsorption onto mild steel occurs by both chemical and physical adsorption, as shown in Table 2.

5 Conclusion

The primary purpose of this research was to explore the potential of an extract derived from *C. nucifera* leaves as a natural inhibitor to improve the resistance of mild steel against corrosion in a seawater environment. A thorough electrochemical analysis, which included techniques such as PP, LPR, assessment of inhibition efficiency, and the application of the Langmuir model to study adsorption isotherms, was conducted to assess the inhibitory characteristics and the underlying adsorption mechanism of the leaf extract.

The findings of this research demonstrate that the incorporation of CNLE in seawater significantly improves the corrosion resistance of mild steel. The PP tests revealed decreased corrosion currents and modified Tafel slopes, indicating a reduction in the corrosion rate and enhanced anodic and cathodic polarization. The LPR results confirmed higher polarization resistance values, confirming the effective inhibition provided by the extract. The inhibition efficiency ultimately reaches its maximum value with increasing extract concentrations, indicating the natural inhibitor's effectiveness. The Langmuir model was used to analyze the adsorption isotherm, which revealed a well-fitted curve showing monolayer adsorption and a significant interaction between the inhibitor and mild steel.

While this study provides promising insights into the corrosion resistance improvement using CNLE, further research is warranted to explore various aspects for practical application. Future work should focus on long-term corrosion testing and environmental factors which investigate the influence of environmental factors, such as temperature and salinity, on the inhibitive properties of the extract in real seawater conditions.

References

1. Xu G et al (2021) Review on corrosion resistance of mild steels in liquid aluminum. J Mater Sci Technol Chin Soc Metals 71:12–22. https://doi.org/10.1016/j.jmst.2020.08.052
2. Al-Amiery AA et al (2023) Exploring the effectiveness of Isatin-Schiff base as an environmentally friendly corrosion inhibitor for mild steel in hydrochloric acid. Lubricants 11:5. https://doi.org/10.3390/lubricants11050211
3. Zhang X et al (2022) Plant extracts as environmentally sustainable corrosion inhibitors I. In: Eco-friendly corrosion inhibitors: principles, designing and applications, pp 263–282. https://doi.org/10.1016/B978-0-323-91176-4.00011-8
4. Al-Moubaraki AH et al (2022) Development of natural plant extracts as sustainable inhibitors for efficient protection of mild steel: experimental and first-principles multi-level computational methods. Materials 15:23. https://doi.org/10.3390/ma15238688
5. Araoyinbo AO et al (2018) Corrosion protection of mild steel in sea water using chemical inhibitor. In: IOP conference series: materials science and engineering. Institute of Physics Publishing. https://doi.org/10.1088/1757-899X/343/1/012012
6. Bahremand F et al (2023) Sustainable development of an effective anti-corrosion film over the St12-steel surface against seawater attacks using Ce(III) ions/tri-sodium phosphate anions. Sci Rep 13(1):12169. https://doi.org/10.1038/s41598-023-38540-9

7. Berrissoul A et al (2020) Evaluation of Lavandula mairei extract as green inhibitor for mild steel corrosion in 1 M HCl solution. Experimental and theoretical approach. J Mol Liq 313. https://doi.org/10.1016/j.molliq.2020.113493
8. Campos MDRS et al (2022) Cathodic protection of mild steel using aluminium-based alloys. Materials 15:4. https://doi.org/10.3390/ma15041301
9. Chen S et al (2020) Corrosion inhibition performance of coconut leaf extract as a green corrosion inhibitor for X65 steel in hydrochloric acid solution. Int J Electrochem Sci 15(1):1–15. https://doi.org/10.20964/2020.01.39
10. Chung IM et al (2020) Inhibition of mild steel corrosion using *Magnolia kobus* extract in sulphuric acid medium. Mater Today Commun 25. https://doi.org/10.1016/j.mtcomm.2020.101687
11. El-Katori EE, Al-Mhyawi S (2019) Assessment of the Bassia muricata extract as a green corrosion inhibitor for aluminum in acidic solution. https://doi.org/10.1080/17518253.2019.1569728
12. Goyal M et al (2018) Organic corrosion inhibitors for industrial cleaning of ferrous and non-ferrous metals in acidic solutions: a review. https://doi.org/10.1016/j.molliq.2018.02.045
13. Guedes D et al (2021) Proanthocyanidins with corrosion inhibition activity for AISI 1020 carbon steel under neutral pH conditions of coconut (*Cocos nucifera* L.) husk fibers. ACS Omega 6(10):6893–6901. https://doi.org/10.1021/acsomega.0c06104
14. Haldhar R et al (2021) Investigation of plant waste as a renewable biomass source to develop efficient, economical and eco-friendly corrosion inhibitor. J Mol Liq 335. https://doi.org/10.1016/j.molliq.2021.116184
15. Hasanah U, Dewi EPS, Santoso U, Supriyadi S (2021) Physical and chemical characteristic of young coconut leaves (*Cocos nucifera* L.) as traditional packaging. Indonesian Food Nutr Prog 17(2):33. https://doi.org/10.22146/ifnp.46769
16. Liu Q et al (2020) A novel green reinforcement corrosion inhibitor extracted from waste *Platanus acerifolia* leaves. Constr Build Mater 260. https://doi.org/10.1016/j.conbuildmat.2020.119695
17. Magga R et al (2021) Coconut water as organic inhibitor on corrosion of low carbon steel in 3.5% NaCl solution. INTEK: Jurnal Penelitian 7(2):145. https://doi.org/10.31963/intek.v7i2.2690
18. Marzorati S et al (2019) Green corrosion inhibitors from natural sources and biomass wastes. Molecules 24:1. https://doi.org/10.3390/molecules24010048
19. Medupin RO et al (2023) Sustainable approach for corrosion control in mild steel using plant-based inhibitors: a review. Mater Today Sustain 22:100373. https://doi.org/10.1016/J.MTSUST.2023.100373
20. Miralrio A, Vázquez AE (2020) Plant extracts as green corrosion inhibitors for different metal surfaces and corrosive media: a review. https://doi.org/10.3390/PR8080942
21. Qiang Y et al (2018) Evaluation of Ginkgo leaf extract as an eco-friendly corrosion inhibitor of X70 steel in HCl solution. Corros Sci 133:6–16. https://doi.org/10.1016/j.corsci.2018.01.008
22. Salinas-Solano G et al (2018) Development and evaluation of a green corrosion inhibitor based on rice bran oil obtained from agro-industrial waste. Ind Crops Prod 119:111–124. https://doi.org/10.1016/j.indcrop.2018.04.009
23. Sanni O et al (2022) Agro-industrial wastes as corrosion inhibitor for 2024–T3 aluminum alloy in hydrochloric acid medium. Res Eng 16:100676. https://doi.org/10.1016/J.RINENG.2022.100676
24. Subekti N et al (2020) Development of environmental friendly corrosion inhibitor from the extract of areca flower for mild steel in acidic media. Eastern-Eur J Enterp Technol 2(6):104, 34–45. https://doi.org/10.15587/1729-4061.2020.197875
25. Sun X et al (2022) Cabbage extract as an eco-friendly corrosion inhibitor for X70 steel in hydrochloric acid medium. J Mol Liq 362:119733. https://doi.org/10.1016/j.molliq.2022.119733
26. Thakur A et al (2022) Plant extracts as environmentally sustainable corrosion inhibitors II. In: Eco-friendly corrosion inhibitors: principles, designing and applications. https://doi.org/10.1016/B978-0-323-91176-4.00017-9

27. Umoren SA et al (2018) Comparative studies on the corrosion inhibition efficacy of ethanolic extracts of date palm leaves and seeds on carbon steel corrosion in 15% HCl solution. J Adhes Sci Technol 32(17):1934–1951. https://doi.org/10.1080/01694243.2018.1455797
28. Wan S, Chen H, Zhang T, Liao B, Guo X (2021) Anti-corrosion mechanism of parsley extract and synergistic iodide as novel corrosion inhibitors for carbon steel-Q235 in acidic medium by electrochemical, XPS and DFT methods. Front BioengBiotechnol 9. https://doi.org/10.3389/fbioe.2021.815953
29. Wan S et al (2021) Anticorrosive reinforcement of waterborne epoxy coating on Q235 steel using NZ/BNNS nanocomposites. Prog Org Coat 159. https://doi.org/10.1016/j.porgcoat.2021.106410

A Comprehensive Analysis of: A Systematic Review

Dian Lestari, Fairuz Iqbal Maulana, Agung Purnomo, and Puput Dani Prasetyo Adi

Abstract This systematic review explores the rising use of virtual clinics, including telemedicine technologies like virtual visits and virtual clinics, in maternity care. These technologies enable continuous monitoring of expectant mothers, offering valuable and satisfactory services, especially in the post-COVID-19 pandemic era. The study aims to comprehensively analyze the applications, effectiveness, benefits, challenges, and overall impact of virtual clinics in gynecological and pregnancy care. A systematic review search encompassing the period from 2019 to 2023 was performed on Scopus and PubMed databases. Additional manual searches on reference lists were conducted for comprehensive coverage. Eight articles were included, comprising various study designs, such as experimental, cohort, and cross-sectional studies. Pregnant women, parents, and healthcare practitioners were the target users. Virtual clinics were utilized for virtual visits, consultations, monitoring, follow-up, and home care, resulting in highly satisfactory patient experiences. However, challenges, including poor internet connection, were identified. The reviewed studies reveal promising outcomes in patient and provider satisfaction. Virtual clinics have the potential to revolutionize maternity care by providing accessible, efficient, and personalized healthcare services. Despite technical challenges, the integration of telemedicine solutions, particularly virtual clinics, is expected to play a growingly significant role in obstetric care. These insights are valuable for healthcare providers, policymakers, and researchers, leading to improved maternal care outcomes through the adoption of virtual clinics.

D. Lestari (✉)
Faculty of Medicine, Universitas Airlangga, Surabaya 60132, Indonesia
e-mail: dian.lestari-2020@fk.unair.ac.id

F. I. Maulana
Computer Science Department, Bina Nusantara University, Jakarta 11480, Indonesia

A. Purnomo
Entrepreneurship Department, BINUS Business School Undergraduate Program, Bina Nusantara University, Jakarta 11480, Indonesia

P. D. P. Adi
Telecommunication Research Center, National Research and Innovation Agency, Jakarta 11480, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_30

Keywords Virtual clinic · Maternity care · Systematic review

1 Introduction

Maternity care is an important aspect of healthcare that aims to provide women with the support and care they need during pregnancy, labor, and the postpartum period [1–3]. The traditional model of maternity care involves regular visits to a health facility, where women receive prenatal care, give birth, and receive postpartum care. However, this model of care has several limitations, including limited access to care, long waiting times, and high costs [4–6]. In recent years, virtual clinics have emerged as a promising alternative to the traditional maternity care model. Virtual clinics use technology to provide remote care to women, so that they can receive care from their homes [7, 8].

Virtual clinics have emerged as a new healthcare approach, garnering increased attention in recent years owing to their ability to enhance healthcare accessibility, mitigate expenses, and address the exigencies of the COVID-19 epidemic [9, 10]. Virtual clinics utilize various technologies, including video conferencing, smartphone applications, and remote monitoring equipment, to deliver healthcare services to patients in a remote manner. The aforementioned strategy offers numerous advantages, encompassing enhanced convenience, diminished trip duration and expenses, and enhanced healthcare accessibility for people residing in rural or geographically isolated regions [9].

Virtual clinics have been implemented throughout several healthcare domains, encompassing mental health, chronic illness management, and primary care. However, its use in maternity care is relatively new. Virtual clinics in maternity care aim to give women access to high-quality care from the comfort of their homes. This approach has several potential benefits, including increased patient satisfaction, reduced costs, and improved health outcomes [11, 12].

Previuos study about virtual clinics in pregnancy and postpartum healthcare underscored virtual prenatal care led to significant reductions in commuting time and work absenteeism, decreases in waiting times at clinics and the rate of missed appointments, minimized the potential for exposure during a pandemic, and promoted greater individual responsibility for healthcare. However, this study did not delve into the examination of how the implementation of virtual clinic modalities and infrastructure affects outcomes and the benefits they provide. As such, these investigations imply that there might be aspects of virtual clinics in maternity care that haven't been thoroughly investigated [10, 13, 14]

Despite the potential benefits of virtual clinics in maternity care, there is a lack of comprehensive analysis of their applications, features, and infrastructure. Such an analysis is necessary to identify the best practices and challenges of virtual clinics in maternity care and to inform the development of effective virtual clinics that meet the needs of pregnant women and new mothers. Therefore, this study aims to conduct a systematic review of virtual clinics in maternity care to identify their applications,

features, infrastructure, and how the benefits and challenges. The results of this study have the potential to contribute valuable insights to the advancement of virtual clinics in the field of maternity care, with the aim of enhancing accessibility to healthcare services and improving patient outcomes.

2 Method

The PRISMA framework, also known as Preferred Reporting Items for Systematic Reviews and Meta-Analyses, was utilized in this comprehensive examination of the existing literature. The PRISMA framework functions as a valuable tool and guide for the assessment of systematic reviews and/or meta-analyses. The checklist encompasses a set of guidelines that provide direction on the fundamental components that have to be incorporated and comprehensively expounded upon in a systematic review and meta-analysis [15–17]. The PRISMA framework utilized in this study is elaborated upon in Fig. 2.

2.1 Search Strategy

A systematic search was performed on the digital databases of Scopus and Pubmed to identify relevant literature pertaining to the utilization of virtual clinics in the context of maternal care. The scope of the inquiry was restricted to scholarly publications composed in the English language and disseminated between the years 2019 and 2023. The investigation employed precise terms, including "virtual clinic," "telehealth," "telemedicine," "maternity," and "perinatal." MESH phrases were additionally employed in the Pubmed database to enhance the precision of the search. At the outset, a total of 128 articles were identified, however, 19 articles were subsequently excluded from the analysis on account of duplication.

2.2 Eligibility Criteria

The present study employed pre-established criteria to ascertain the literature that would be incorporated or rejected for the purpose of the investigation. The inclusion criteria consisted of studies that examined the use of virtual clinics in maternity care, evaluated their effectiveness, acceptability, satisfaction, or economic implications, employed quantitative, qualitative, or mixed-methods research designs, and were published in peer-reviewed journals. In contrast, the exclusion criteria consisted of systematic reviews, meta-analyses, and bibliometric reviews. Additionally, studies examining virtual clinics in healthcare settings other than maternity care

were excluded. Furthermore, studies that did not present primary data or original research findings, as well as studies not published in English, were also excluded.

2.3 Article Screening

The investigation was initiated by acquiring pertinent material from digital databases such as Pubmed and Scopus, which was subsequently organized and stored in a Mendeley library. In order to uphold the standard and uniformity of the literature, any duplicate articles were eliminated, resulting in a total of 128 articles that met the criteria for subsequent evaluation. Upon careful examination of the titles, a total of 41 papers were deemed unsuitable and hence deleted from the study. Furthermore, an additional 19 publications were eliminated due to their failure to match the specified exclusion criteria. Following the completion of the initial phases, a total of eight articles were identified and then incorporated into the systematic review.

2.4 Bias Assessment

The systematic review included a total of eight papers, which were selected for inclusion based on predetermined criteria. To ensure the quality and reliability of the included studies, a bias assessment was conducted using the CASP (Critical Appraisal Skills Programme) checklist. The present checklist comprises a set of ten inquiries that have been specifically formulated to assess the efficacy of clinical prediction rules. The quality of each study was assessed using the CASP critical score, which allocates 2 points for fully satisfying criteria, 1 point for partially satisfying criteria, and 0 points for criteria that are not applicable, not satisfied, or not specified. Following the evaluation, the studies were categorized according to their level of quality. A cumulative score of 22 signifies a high level of quality, whilst a score falling between the range of 16 to 21 indicates a moderate level of quality. On the contrary, a score of $\leq$15 signifies a diminished level of quality [18]. The purpose of this assessment is to ascertain the reliability and adequacy of the selected papers for inclusion in the systematic review [19].

3 Results

3.1 Study Characteristic

The methodology used for article categorization followed the PRISMA approach, where articles were classified based on predefined criteria that included title, abstract, and full content. Initially, a corpus of 128 articles was identified, and after eradication of duplication and application of exclusion parameters, a final selection of 8 articles was deemed suitable for systematic review. The manuscripts examined showed a diverse range of research typologies, consisting of 2 cohort studies (25%), 3 experimental studies (37.5%), and 3 cross-sectional investigations (37.5%). Geographically, the study populations were mainly from Europe, America, and Asia. After bias assessment through the use of CASP, findings showed that the articles exhibited a moderate level of quality, with no examples categorized as poor quality, according to the evaluation criteria. The total composite score for the assessment ranged between 18 and 20. Eight selected articles underwent a thorough review. The reviewed variables included study type, target group, virtual clinic objectives, modalities, study outcomes, benefits, and challenges. The review results are elaborated in Table 1.

The research on virtual clinics in maternity care has been steadily increasing over the years. The increase in research activity within this field over the previous decade is clearly demonstrated by the data obtained from the Scopus database, as depicted in Fig. 1. The trend has experienced a notable surge, particularly since 2019, with this accelerated growth being attributable to the COVID-19 pandemic. The global epidemic necessitated modifications in healthcare policies and imposed restrictions on in-person healthcare provisions. Consequently, the implementation of virtual clinics emerged as a pragmatic resolution, thereby fostering further scholarly investigation and inquiry within this domain. The year 2019 marked a turning point as the pandemic accelerated the adoption of virtual clinics for maternity care. This shift led to more research focusing on the effectiveness and patient satisfaction of virtual maternity care, transforming how maternity care is approached and delivered (Fig. 2).

3.2 Virtual Clinics in Maternity Clinic

4 Discussion

This review was undertaken due to the increasing adoption of telemedicine for pregnancy care, postpartum support, and newborn care. The studies we incorporated centered around the development of virtual systems for such care, encompassing a

Table 1 Applications, modalities, study outcomes, benefits, and challenges in virtual clinic of maternity care

Document	Type of study	Target group	Purpose of virtual clinic	Modalities	Study outcomes	Benefits	Challenges
[20]	Experimental (2022)	• 95 pregnant woman • 96 post-partum woman	Ameliorating post-partum depression	• Mommoodbooster (m-Health) • Routine care	Decreased severity of depression, anxiety, stress, and automatic thoughts, and increased behavioral activation and self-efficacy	Effective for depression treatment especially when combined with universal depression screening and referral	Assessment over a long period of time and examination of how long the treatment effect lasts
[21]	Retrospective study (2020)	498 patients with hypertensive disorders of pregnancy	Diagnosis of HDP and severity	• Routine prenatal visits • Telehealth visits	• Using telehealth is secure for incorporating prenatal care and monitoring high blood pressure disorders during pregnancy • Telehealth did not cause any delays or worsened diagnoses of high blood pressure disorders during pregnancy	N/A	Telehealth approaches varied among patients, with differences observed in aspects such as the choice between video and audio consultations

(continued)

Table 1 (continued)

Document	Type of study	Target group	Purpose of virtual clinic	Modalities	Study outcomes	Benefits	Challenges
[22]	Prospective cohort study (2019–2020)	174 average risk pregnant women	Prenatal monitoring	• Phone visits • Online portal visits • Doppler • Blood pressure cuff	Telehealth experiences were excellent and just as good as in-person visits	• Easy and wide accessibility • Saved time • Convenient	• Weak audiovisual connection • Technical issues on virtual platforms • Lack of real-time customer support • Enhancing Doppler/blood pressure cuff instructions
[23]	Experimental (2018–2019)	• 42 parents of children less than 2 years of age • 128 expectant parents	Breastfeeding education intervention for fathers and mothers in a low resource setting	SMS text message-based health education	• Survey participants were very satisfied and willing to keep receiving SMS text messages • SMS text messages effectively boosted exclusive breastfeeding rates during the initial 3 months	• Convenient • Parents feel supported in the breastfeeding process	N/A

(continued)

Table 1 (continued)

Document	Type of study	Target group	Purpose of virtual clinic	Modalities	Study outcomes	Benefits	Challenges
[24]	Quasy experiment (2020)	4501 pregnant women	Pregnancy and postpartum depression health education and ANC visit reminder	SMS text based comprehensive intervention	• The intervention group showed significantly lower depression screening results compared to the control group • Help reduce depressive symptoms and lower the risk of perinatal depression	• Convenient • Low-cost	• Lack of internet access • Long-term follow-up
[25]	Cross sectional study (2019)	108 pregnant women	Pregnancy self-monitor	Digital health application	Pregnant women feel the need for better pregnancy tools for comprehensive care from before to after delivery that can be linked to the healthcare system	• Consistent and convenient tracking of changes in body shape over time • Broader range of comprehensive data concerning pregnancy-related body transformations	Overinterpreting results and wrongly linking to disease risk

(continued)

Table 1 (continued)

Document	Type of study	Target group	Purpose of virtual clinic	Modalities	Study outcomes	Benefits	Challenges
[26]	Cross sectional study (2019)	• 50 midwives • 29 gynaecologists	Monitoring high-risk pregnancy	• Mobile-based telehealth • Web-based telehealth	Mobile-based technologies was found to be more practical than web-based methods for managing high-risk pregnancies	Mobile-based applications are more user-friendly and familiar to navigate	Security of patient confidentiality and privacy
[27]	Mixed-methods cross-sectional study (2021)	31 healthcare practitioners	Services for discussing and assisting with perinatal issues extended to rural and distant areas	Telehealth psychiatry consultation (swops)	Increased knowledge and confidence in women with moderate and mild mental health conditions	Wider range education topic	• Poor awareness of service and referral pathways • The service was time-consuming

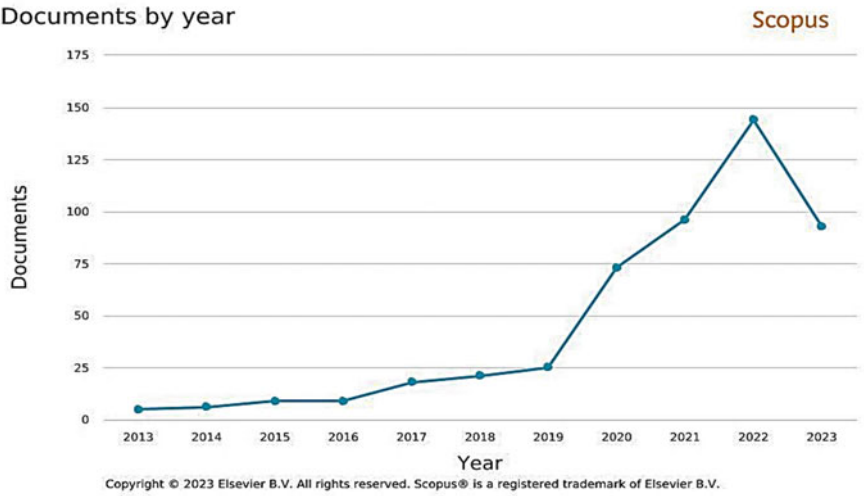


Fig. 1 Virtual clinic in maternity care research trend over decade

blend of in-person and online services. In general, there appeared to be no significant distinction in terms of patient satisfaction and the outcomes of care provision between live and virtual modalities.

Virtual care methodologies have exhibited positive outcomes in the management and monitoring of high-risk pregnancies. Following the COVID-19 pandemic, medical guidelines advocated for the utilization of remote care visits through "telehealth" technology to mitigate viral transmission [28–30]. This led to a reduction in physical visits, resulting in time and cost savings while enhancing satisfaction levels for both patients and providers. However, the adoption of remote care also introduces technical and infrastructural challenges [8, 10].

The implementation of telemedicine technologies can curtail the frequency of in-person prenatal visits. Despite medical recommendations suggesting around 12–14 care appointments [31, 32], some studies indicate that this number can be scaled down to as few as 9 visits, or even fewer, through the substitution of physician visits with nurse or midwife consultations via virtual platforms [33]. The timeline for virtual visits varies among studies, but the general trend is to retain some face-to-face interactions, especially during critical pregnancy milestones, while virtual appointments become more prevalent as the pregnancy advances due to heightened risk factors. Although the rate of missed appointments remains consistent between physical and virtual visits, the motivations behind these absences might diverge. Concerns about virus transmission might influence in-person attendance, especially post the COVID-19 pandemic, whereas technical obstacles and accessibility issues could impact virtual engagements [34–36]

To ensure the efficacy of virtual appointments, appropriate equipment and infrastructure are imperative [37, 38]. Healthcare providers need access to computers,

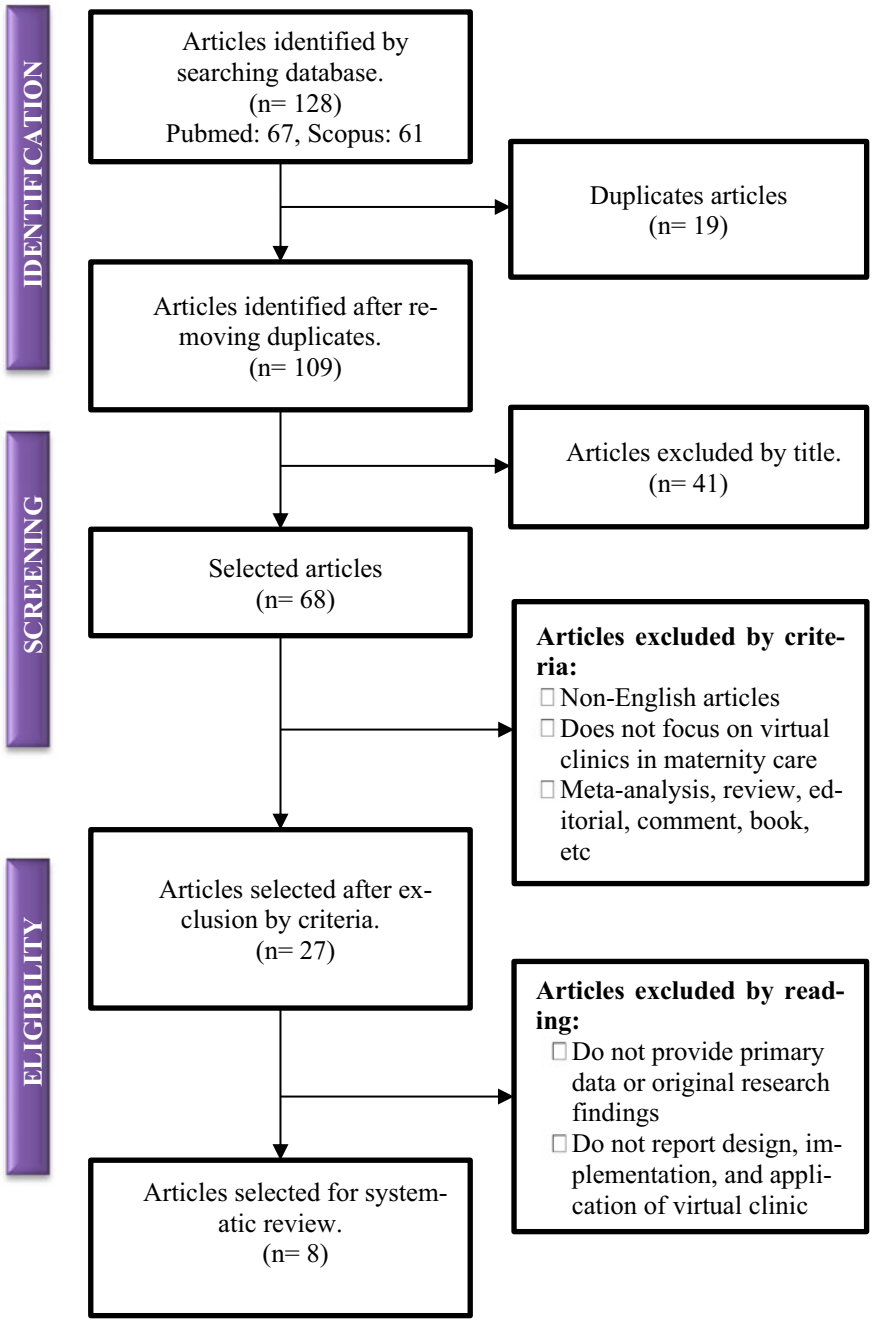


Fig. 2 PRISMA flowchart for selecting articles

tablets, or smartphones equipped with video call capabilities and reliable internet connectivity [39]. Likewise, patients require similar devices for their participation in virtual visits, educational sessions, and the assessment of health progress. Portable tools such as blood pressure monitors, handheld pulse oximeters, Doppler devices, and cardiotocograph devices are prerequisites for remote monitoring [40]. During virtual consultations, pregnant individuals can assess their weight, blood pressure, fetal heart rate, uterine contractions, and uterine fundal height with the guidance of healthcare professionals [41, 42]. However, certain tests, like non-stress tests, might be infeasible at home due to limitations in commercially available equipment. Medical workers save data generated by this technology and symptom-related information in electronic medical records or cloud-based systems. Although the cost of equipment and its funding sources are not underscored in this study, patients might either purchase the equipment themselves, receive it from the hospital, or secure funding from alternate sources [43].

Obstacles to the uptake of telehealth can be divided into two categories. Firstly, a secure and stable connection between patients and healthcare providers is paramount [44]. A weak connection can disrupt communication, leading to frustration for all parties involved. The quality of the connection plays a pivotal role in overall patient contentment. Several issues, including technical glitches, inadequate or absent internet access, and a shortage of suitable equipment, can impede the seamless operation of telehealth services [45, 46]. Additionally, the initial expenses can pose a significant barrier to implementation. Other impediments encompass a lack of technological literacy, insufficient trust, conservative attitudes, and the perception that virtual visits lack emotional efficacy [47, 48]. Some studies have also highlighted patients' challenges in adapting to video technology [49]. Moreover, tools such as Doppler devices and non-stress tests might be hard to access, potentially limiting the assessment of fetal movements. These challenges might also vary based on specific medical divisions, like emergency departments, NICUs, and PICUs [50]. While perspectives regarding the benefits, obstacles, and overall impact of telemedicine vary between different medical departments, issues such as poor internet connectivity, mobile app difficulties, and challenges associated with desktop usage are commonly reported [49].

5 Findings

The findings of this systematic review indicate that virtual prenatal care holds significant promise, addressing several crucial aspects of maternal care. It has the potential to break down access barriers, enabling women to engage actively in their care while proving both cost-effective and adaptable based on resource availability and patient preferences. Moreover, the feasibility, acceptance, and effectiveness of virtual

prenatal care in enhancing maternal and neonatal outcomes were evident, accompanied by heightened patient satisfaction, reduced travel expenses and time commitments. Impressively, this mode of care minimizes in-person visits without compromising quality, emerging as a valuable tool for monitoring high-risk pregnancies and delivering timely interventions. However, the studies underscored the necessity of equitable accessibility, ensuring women from all backgrounds benefit, and emphasized further research to uncover the long-term implications of virtual prenatal care on maternal and neonatal wellbeing.

6 Conclusion

The present systematic review examined a total of eight scholarly studies that documented the utilization and outcomes of virtual clinics in the context of pregnancy and postpartum healthcare. The research encompassed a range of study designs, including experimental, cohort, and cross-sectional investigations. The participants of interest were pregnant women, parents, and healthcare practitioners. Virtual clinics have been utilized for various purposes, including remote visits, consultations, monitoring, follow-up, and the provision of home care through virtual means. The studies that were analyzed demonstrated encouraging results in terms of patient and provider satisfaction. The implementation of virtual clinics in the context of prenatal care has promise in mitigating obstacles to accessing healthcare services, particularly those arising from issues related to transportation, employment, and childcare. The study posits that telehealth is poised to emerge as a prevalent modality for delivering healthcare services to expectant mothers. The present study offers a thorough examination of virtual clinics in the context of maternity care, shedding light on their various uses, characteristics, and infrastructure within the field of gynecological and prenatal care.

It is hoped that future research will expand to assess the acceptability and feasibility of virtual clinics among various populations, including high-risk patients, low-income groups, and marginalized communities.

References

1. Antenatal care coverage—at least four visits. https://www.who.int/data/gho/indicator-metadata-registry/imr-details/80. Last accessed 05 May 2020
2. WHO (2011) Integrated management of pregnancy and childbirth (IMPAC). WHO
3. de Masi S, Bucagu M, Tunçalp Ö, Peña-Rosas JP, Lawrie T, Oladapo OT, Gülmezoglu M (2017) Integrated person-centered health care for all women during pregnancy: implementing world health organization recommendations on antenatal care for a positive pregnancy experience. Glob Health Sci Pract 5:197–201. https://doi.org/10.9745/GHSP-D-17-00141
4. Nguyen E, Engle G, Subramanian S, Fryer K (2021) Telemedicine for prenatal care: a systematic review. medRxiv. 2021.05.14.21257232
5. Shaw D, Guise J-M, Shah N, Gemzell-Danielsson K, Joseph KS, Levy B, Wong F, Woodd S, Main EK (2016) Drivers of maternity care in high-income countries: can health systems support

woman-centred care? Lancet 388:2282–2295. https://doi.org/10.1016/S0140-6736(16)31527-6
6. McLeish J, Redshaw M (2019) Maternity experiences of mothers with multiple disadvantages in England: a qualitative study. Women Birth 32:178–184. https://doi.org/10.1016/j.wombi.2018.05.009
7. Butler Tobah YS, LeBlanc A, Branda ME, Inselman JW, Morris MA, Ridgeway JL, Finnie DM, Theiler R, Torbenson VE, Brodrick EM, Meylor de Mooij M, Gostout B, Famuyide A (2019) Randomized comparison of a reduced-visit prenatal care model enhanced with remote monitoring. Am J Obstet Gynecol 221:638.e1–638.e8. https://doi.org/10.1016/j.ajog.2019.06.034
8. Cantor AG, Jungbauer RM, Totten AM, Tilden EL, Holmes R, Ahmed A, Wagner J, Hermesch AC, McDonagh MS (2022) Telehealth strategies for the delivery of maternal health care a rapid review. Ann Intern Med 175:1285–1297. https://doi.org/10.7326/M22-0737
9. Craighead CG, Collart C, Frankel R, Rose S, Misra-Hebert AD, Tucker Edmonds B, Michie M, Chien E, Coleridge M, Goje O, Ranzini AC, Farrell RM (2022) Impact of telehealth on the delivery of prenatal care during the COVID-19 pandemic: mixed methods study of the barriers and opportunities to improve health care communication in discussions about pregnancy and prenatal genetic testing. JMIR Form Res 6:e38821. https://doi.org/10.2196/38821
10. Hofmann G, Hampanda K, Harrison MS, Fasano M, Nacht A, Yeoman M (2022) Virtual prenatal and postpartum care acceptability among maternity care providers. Matern Child Health J 26:1401–1408. https://doi.org/10.1007/s10995-022-03412-7
11. Øvretveit J (2021) Innovations in self care and close care made during COVID 19 pandemic: a narrative review. Int J Health Gov 26:88–99. https://doi.org/10.1108/IJHG-02-2021-0007
12. Galle A, Semaan A, Huysmans E, Audet C, Asefa A, Delvaux T, Afolabi BB, El Ayadi AM, Benova L (2021) A double-edged sword-telemedicine for maternal care during COVID-19: findings from a global mixed-methods study of healthcare providers. BMJ Glob Health 6. https://doi.org/10.1136/bmjgh-2020-004575
13. Ghimire S, Martinez S, Hartvigsen G, Gerdes M (2023) Virtual prenatal care: a systematic review of pregnant women's and healthcare professionals' experiences, needs, and preferences for quality care. Int J Med Inform 170:104964. https://doi.org/10.1016/j.ijmedinf.2022.104964
14. Thirugnanasundralingam K, Davies-Tuck M, Rolnik DL, Reddy M, Mol BW, Hodges R, Palmer KR (2023) Effect of telehealth-integrated antenatal care on pregnancy outcomes in Australia: an interrupted time-series analysis. Lancet Digit Health 5:e798–e811. https://doi.org/10.1016/S2589-7500(23)00151-6
15. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, Moher D, Peters MDJ, Horsley T, Weeks L, Hempel S, Akl EA, Chang C, McGowan J, Stewart L, Hartling L, Aldcroft A, Wilson MG, Garritty C, Lewin S, Godfrey CM, MacDonald MT, Langlois EV, Soares-Weiser K, Moriarty J, Clifford T, Tunçalp Ö, Straus SE (2018) PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 169:467–473. https://doi.org/10.7326/M18-0850
16. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372. https://doi.org/10.1136/bmj.n71
17. Kamioka H (2019) Preferred reporting items for systematic review and meta-analysis protocols (prisma-p) 2015 statement. Jpn Pharmacol Ther 47:1177–1185
18. Zeng X, Zhang Y, Kwong JSW, Zhang C, Li S, Sun F, Niu Y, Du L (2015) The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. J Evid Based Med 8:2–10. https://doi.org/10.1111/jebm.12141
19. Harrison JK, Reid J, Quinn TJ, Shenkin SD (2017) Using quality assessment tools to critically appraise ageing research: a guide for clinicians. Age Ageing 46:359–365. https://doi.org/10.1093/ageing/afw223

20. Danaher BG, Seeley JR, Silver RK, Tyler MS, Kim JJ, La Porte LM, Cleveland E, Smith DR, Milgrom J, Gau JM (2023) Trial of a patient-directed eHealth program to ameliorate perinatal depression: the MomMoodBooster2 practical effectiveness study. Am J Obstet Gynecol 228(453):e1-453.e10. https://doi.org/10.1016/j.ajog.2022.09.027
21. Abelman SH, Svetec S, Felder L, Boelig RC (2023) Impact of telehealth implementation on diagnosis of hypertensive disorders of pregnancy. Am J Obstet Gynecol MFM 5:101043. https://doi.org/10.1016/j.ajogmf.2023.101043
22. Bruno B, Mercer MB, Hizlan S, Peskin J, Ford PJ, Farrell RM, Rose SL (2023) Virtual prenatal visits associated with high measures of patient experience and satisfaction among average-risk patients: a prospective cohort study. BMC Pregnancy Childbirth 23:1–11. https://doi.org/10.1186/s12884-023-05421-y
23. Gebremariam KT, Mulugeta A, Gallegos D (2023) Co-design and implementation of a mHealth intervention targeting fathers and mothers to improve breastfeeding. BMC Med Inform Decis Mak 23:1–7. https://doi.org/10.1186/s12911-023-02125-3
24. Gong M, Zhang S, Xi C, Luo M, Wang T, Wang Y, Wang S, Guo L, Lu C (2021) Comprehensive intervention during pregnancy based on short message service to prevent or alleviate depression in pregnant women: a quasi-experimental study. Early Interv Psychiatry 15:352–359. https://doi.org/10.1111/eip.12953
25. Lazarevic N, Pizzuti C, Rosic G, Bœhm C, Williams K, Caillaud C (2023) A mixed-methods study exploring women's perceptions and recommendations for a pregnancy app with monitoring tools. NPJ Digit Med 6. https://doi.org/10.1038/s41746-023-00792-0
26. Ayatollahi H, Ghalandar Abadi M, Hemmat M (2019) Web and mobile-based technologies for monitoring high-risk pregnancies? BMJ Health Care Informatics 26:1–5. https://doi.org/10.1136/bmjhci-2019-000025
27. Cibralic S, Fay-Stammbach T, Tucker D, Song D, Eapen V (2023) A mixed-methods evaluation of a state-wide outreach perinatal mental health service. BMC Pregnancy Childbirth 23:1–7. https://doi.org/10.1186/s12884-022-05229-2
28. Denicola N (2020) Implementing telehealth in practice: ACOG committee opinion summary, number 798. Obstet Gynecol 135:493–494
29. Vernon MM, Yang FM (2022) Implementing a self-monitoring application during pregnancy and postpartum for rural and underserved women: a qualitative needs assessment study. PLoS ONE 17:e0270190. https://doi.org/10.1371/journal.pone.0270190
30. Brown HL, DeNicola N (2020) Telehealth in maternity care. Obstet Gynecol Clin North Am 47:497–502. https://doi.org/10.1016/j.ogc.2020.05.003
31. ACOG (2022) Guidelines for perinatal care. https://doi.org/10.1007/978-3-030-90642-9_3
32. NHS (2021) Antenatal care: timing of first antenatal appointment. https://doi.org/10.1136/bmj.1.3570.1058
33. ACOG (2020) Alternate or reduced prenatal care schedules. ACOG Bull 19
34. Konnyu KJ, Danilack VA, Adam GP, Friedman Peahl A, Cao W, Balk EM (2023) Changes to prenatal care visit frequency and telehealth: a systematic review of qualitative evidence. Obstet Gynecol 141:299–323. https://doi.org/10.1097/AOG.0000000000005046
35. Shmerling A, Hoss M, Malam N, Staton EW, Lyon C (2022) Prenatal care via telehealth. Prim Care 49:609–619. https://doi.org/10.1016/j.pop.2022.05.002
36. Morgan A, Goodman D, Vinagolu-Baur J, Cass I (2022) Prenatal telemedicine during COVID-19: patterns of use and barriers to access. JAMIA Open 5:ooab116. https://doi.org/10.1093/jamiaopen/ooab116
37. Stanley AY, Wallace JB (2022) Telehealth to improve perinatal care access. MCN Am J Matern Nurs 47:281–287. https://doi.org/10.1097/NMC.0000000000000841
38. Hawkins SS (2023) Telehealth in the prenatal and postpartum periods. JOGNN—J Obstet Gynecol Neonatal Nurs 52:264–275. https://doi.org/10.1016/j.jogn.2023.05.113
39. Aziz A, Zork N, Aubey JJ, Baptiste CD, D'alton ME, Emeruwa UN, Fuchs KM, Goffman D, Gyamfi-Bannerman C, Haythe JH, Lasala AP, Madden N, Miller EC, Miller RS, Monk C, Moroz L, Ona S, Ring LE, Sheen JJ, Spiegel ES, Simpson LL, Yates HS, Friedman AM (2020) Telehealth for high-risk pregnancies in the setting of the COVID-19 pandemic. Am J Perinatol 37:800–808. https://doi.org/10.1055/s-0040-1712121

40. Atkinson J, Hastie R, Walker S, Lindquist A, Tong S (2023) Telehealth in antenatal care: recent insights and advances. BMC Med 21:1–17. https://doi.org/10.1186/s12916-023-03042-y
41. Jin MX, Kim SY, Miller LJ, Behari G, Correa R (2020) Telemedicine: current impact on the future. https://doi.org/10.7759/cureus.9891
42. Ghimire S, Gerdes M, Hartvigsen G (2021) Virtual prenatal care: a systematic review of pregnant women' and healthcare professionals' experiences, needs, and preferences for quality care Ghimire. Int Platform Reg Syst Rev Meta-Anal 20–21. https://doi.org/10.37766/inplasy2022.9.0070
43. Lukas H, Xu C, Yu Y, Gao W (2020) Emerging telemedicine tools for remote COVID-19 diagnosis, monitoring, and management. ACS Nano 14:16180–16193. https://doi.org/10.1021/acsnano.0c08494
44. Zachrison KS, Boggs KM, Hayden EM, Espinola JA, Camargo CA (2020) Understanding barriers to telemedicine implementation in rural emergency departments. Ann Emerg Med 75:392–399. https://doi.org/10.1016/j.annemergmed.2019.06.026
45. Zhang T, Mosier J, Subbian V (2021) Identifying barriers to and opportunities for telehealth implementation amidst the COVID-19 pandemic by using a human factors approach: a leap into the future of health care delivery? JMIR Hum Factors 8:e24860. https://doi.org/10.2196/24860
46. Hung P, Granger M, Boghossian N, Yu J, Harrison S, Liu J, Campbell BA, Cai B, Liang C, Li X (2023) Dual barriers: examining digital access and travel burdens to hospital maternity care access in the United States, 2020. Milbank Q. https://doi.org/10.1111/1468-0009.12668
47. Parimbelli E, Bottalico B, Losiouk E, Tomasi M, Santosuosso A, Lanzola G, Quaglini S, Bellazzi R (2018) Trusting telemedicine: a discussion on risks, safety, legal implications and liability of involved stakeholders. Int J Med Inform 112:90–98. https://doi.org/10.1016/j.ijmedinf.2018.01.012
48. Haleem A, Javaid M, Singh RP, Suman R (2021) Telemedicine for healthcare: capabilities, features, barriers, and applications. Sens Int 2:100117. https://doi.org/10.1016/j.sintl.2021.100117
49. Ftouni R, AlJardali B, Hamdanieh M, Ftouni L, Salem N (2022) Challenges of telemedicine during the COVID-19 pandemic: a systematic review. BMC Med Inform Decis Mak 22:207. https://doi.org/10.1186/s12911-022-01952-0
50. Cassar MR, Borg D, Camilleri L, Schembri A, Anastasi EA, Buhagiar K, Callus C, Grech M (2021) A novel use of telemedicine during the COVID-19 pandemic. Int J Infect Dis 103:182–187

Enhancing Extractive Summarization in Student Assignments Using BERT and K-Means Clustering



Mamluatul Hani'ah, Vivi Nur Wijayaningrum, and Astrifidha Rahma Amalia

Abstract Evaluation through learning assessments is a fundamental factor in determining student's success in achieving specified competencies. During the essay evaluation, the lecturer needs to check each assignment individually. However, when dealing with long essay answers, extra attention is necessary to extract the key points effectively. As a result, this process takes a lot of time and may potentially lead to carelessness or boredom during the assignment-checking process. The way to overcome this problem is by using Extractive summarization. The Extractive Summarization method summarizes by extracting key points from the student assignments and creating a summary without making any changes to the original text. Currently, Extractive Summarization widely uses deep learning technology, such as Bidirectional Encoder Representations from Transformers (BERT). BERT can effectively recognize contextual information within sentences. Output from BERT is sentence embeddings that can serve as valuable features for clustering. The summary result generated by applying k-means clustering to group sentences that have similarities and then selecting one primary sentence from each cluster to represent the cluster. This research proposed an approach for selecting sentence candidates for each cluster using TF-IDF weighting. The proposed method achieved the best ROUGE score on ROUGE-1 recall with 0.73003. We compare our results with the previous study's BERT k-means approach, which selects sentence candidates from the closest sentences to the centroid for summary selection. The experimental results show that the proposed method achieves slightly better ROUGE scores than the previous study. Furthermore, in terms of execution time comparison, the proposed method has a shorter execution time.

Keywords Student assignments summarization · Deep learning · BERT · k-means clustering · TF-IDF weighting

M. Hani'ah (✉) · V. N. Wijayaningrum · A. R. Amalia
Department of Information Technology, Politeknik Negeri Malang, Malang, Indonesia
e-mail: mamluatulhaniah@polinema.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_31

1 Introduction

Evaluation in the form of learning assessment becomes one of the most important aspects of determining the student's success in achieving the specified competencies. Assessment of learning outcomes can be measured in various ways, both through tests and non-tests [1, 2]. The assessments commonly used by lecturers for evaluations include assignments, quizzes, and exams [3]. Assignments can range from answering essay questions, creating essays, or completing projects.

In assessing the results of student assignments, the lecturer must check each student's assignment. If the assignment is in the form of a long essay, then the lecturer needs more attention to check it to get the main information from it. Consequently, assessing student assignments takes a lot of time, which may lead to unintentional oversights or even boredom during the review process. One way to overcome this problem is to use automatic summarization technology [4]. Automatic Text Summarization (ATS) is a branch of Natural Language Processing (NLP) that can generate core information from a long text [5–7].

Automatic Text Summarization (ATS) has two distinct approaches: the extractive summarization method and the abstractive summarization method [8–10]. The extractive summarization method generates summaries by extracting all key points from the main document, creating a summary without making changes to the original text. It involves selecting sentences that serve as key points in the document and copying these key sentences into the summary [11]. On the other hand, the abstractive summarization method creates summaries by generating new words and sentences that are different from the original document but still represent the original document [12].

Extractive summarization has been widely used in summarizing in the field of education [4, 13, 14]. The study conducted by [4] focuses on the automated summarization of student assignment documents using the Fuzzy C-Means method with sentence weighting features. Sentence weighting is a feature generated through the selection of sentences with the highest weights within a cluster, allowing for rapid extraction of key information from a document. However, the results of this study provide a summary result that cannot be said to be optimal. In recent times, deep learning has made significant progress in the field of Natural Language Processing (NLP), including the domain of summarization tasks. One of the pre-trained network models in deep learning that has shown remarkable progress is BERT [15].

BERT uses a transformer architecture with encoder and decoder layers [16]. A study conducted by [17] performed extractive summarization of lecturer content using BERT. BERT has been proven to be able to understand and process text well. The output from BERT, which is sentence embeddings, serves as feature for k-means clustering to group similar sentences. The sentences closest to the center of each cluster are then selected as cluster representatives, forming the summary output. Even though extractive summarization outcomes are still imperfect, BERT consistently improves summary quality when compared to Text-Rank [17]. BERT is a pre-trained model, so when using it for document summarization in Bahasa, a

BERT model that has been fine-tuned with Indonesian documents is necessary. This study use pre-trained BERT indolem/indobert-base-uncased model by [18].

Clustering in document summarization plays an important role in grouping sentences with similar information. To obtain the main information from a cluster, a selection of core sentences that can represent each cluster is necessary. This sentence selection process can use techniques such as random selection or ranking based on TF-IDF [19]. In this study, document summarization of student assignments is performed using BERT and k-means clustering. BERT is used to generate sentence embeddings, and k-means clustering is used to group sentences based on the sentence embeddings produced by BERT. This research introduces an approach to choosing sentence candidates for each cluster using TF-IDF weighting to generate a final summary.

2 Materials and Method

2.1 Material

This study utilizes student assignment data from the course "Islamic Education" at Politeknik Negeri Malang. The data collected consists of long essays from 62 students with the same assignment topic. This data will subsequently be summarized using the proposed method. To assess the performance of the proposed method, summary documents from experts are required. Therefore, experts (Islamic Education lecturers) need to manually summarize the 62 student assignment documents. The manually summarized documents by the expert will be compared with the summary documents generated by the proposed method during the evaluation phase.

2.2 Method

The research process involves four primary phases: preprocessing, generating sentence embeddings using pretrained IndoBERT, performing sentence clustering using k-means clustering, and finally selecting informative sentences to represent each cluster. These stages are illustrated in Fig. 1.

Preprocessing. The input documents can be in the format of Word or PDF files, which are then extracted using Textract. The extraction results from Textract often include symbols like "\n" for new lines and "\t" for tabs. Therefore, preprocessing is required to clean up these symbols. Apart from removing unused symbols, this stage is also conducted to eliminate unnecessary characters and convert all sentences to lowercase. In document summarization, it's crucial to identify sentences containing key information, so sentences with fewer than three words are removed.

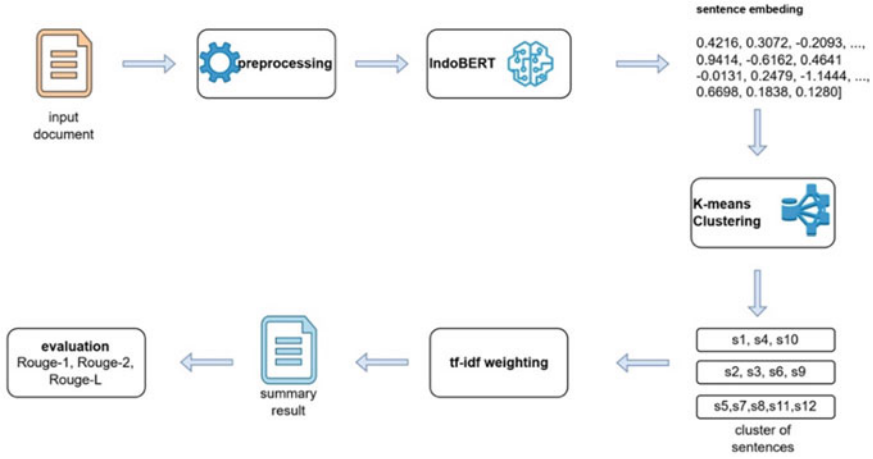


Fig. 1 Research method

IndoBERT. Deep Learning in the field of NLP has provided many extraordinary developments in various problems [20]. Recently, a prominent deep learning model based on the transformer architecture has significantly enhanced performance in NLP [16]. The Transformer Architecture introduced an encoder and decoder setup that can process text bidirectionally, both from left to right and right to left. The encoder is responsible for breaking down input text into vector representations, while the decoder generates text predictions. A widely-used implementation of the Transformer is BERT (Bidirectional Encoder Representations from Transformers). BERT is a pre-train transformer network that can be used in various kinds of NLP tasks, including classification [21], sentiment analysis [22], and document summarization [14, 17]. BERT transforms words into vectors through a three-stage embedding process. (1) **Token Embedding**: BERT incorporates several tokens generated during the tokenization process, including [CLS], [SEP], and [PAD]. The [CLS] token is positioned at the beginning of the text to signify its start. The [SEP] token emerges when there are two sentences in the text or at the end of a sentence, serving as a separator between them. On the other hand, the [PAD] token is employed when the token count from tokenization is less than the predetermined token length in the BERT model [23]. In this study, the token length used varies based on the number of tokens present in the longest sentence. (2) **Segment Embedding**: Segment embedding is employed to differentiate between different segments within the text. This is particularly useful when there are two or more sentences in the text; mark each sentence distinctly. (3) **Position Embedding**: Position embedding provides information about the position of each word within a sentence. [15]

In this study, we employ the pretrained Bert model, IndoBERT [18]. IndoBERT utilizes the BERT-case architecture, featuring 12 transformer layers, each containing 768 hidden units, and 12 attention heads per hidden unit. IndoBERT is a transfer

learning model that has been trained using over 220 million words in the Indonesian language, from various texts including Wikipedia in Indonesian, online news websites, and the Indonesian Web Corpus.

K-means Clustering. Sentence clustering aims to group sentences that share similarities [24]. Essentially, clustering is used to organize structured data [25]. However, the initial data, in the form of text, is unstructured. The output of IndoBERT provides vector representations that transform text into sentence embeddings. BERT's sentence embeddings capture the semantic meaning and context of sentences in numerical vector format, making the IndoBERT output structured and suitable for clustering algorithms. In this study, k-means clustering is employed for sentence grouping. The pseudocode for k-means clustering is illustrated in pseudocode-1:

pseudocode-1:

K-means Clustering:

Input: set of vectors V, number of cluster (k)

Output: cluster of sentences

Process:

1. Select k centroids randomly from sentence embeddings.
2. Calculate the similarity between centroids and non-centroid sentence embeddings.
3. For each row in sentence embeddings, identify cluster members by selecting each row of sentence embedding with the highest similarity to the non-centroid sentence embedding with the centroid
4. Update centroids by calculating the average of cluster members.
5. Repeat steps 2 to 5 until there is no change in cluster members.

TF-IDF Weighting. The summary result is generated by combining sentence candidates from each cluster. Each cluster will possess a sentence candidate from within its own cluster. Thus, the number of sentences in the summary is equivalent to the number of clusters. The selection of sentence candidates from each cluster is based on sentences with the highest Term Frequency—Inverse Document Frequency (TF-IDF) values. TF-IDF is a word weighting technique that surpasses word frequency [26]. Term Frequency refers to how often a word appears in a document. Inverse Document Frequency indicates how common a word is across documents. In TF-IDF, words that appear rarely in documents are considered more significant than words that appear very frequently. The TF-IDF formula can be seen in (1),(2),(3) [27].

$$tf(w) = \frac{number\ of\ times\ w\ appears\ in\ D}{Sum\ of\ terms\ in\ D} \tag{1}$$

$$idf(w) = \log \frac{total\ number\ of\ D}{number\ of\ D\ contain\ w} \tag{2}$$

$$tf - idf(w) = tf(w) * idf(w) \tag{3}$$

To calculate TF-IDF, calculate TF (term frequency) value, which is the number of times "w" appears in document D divided by the total number of terms in document D. Furthermore, the IDF (inverse document frequency) value of the word "w" is calculated as the logarithm of the total number of documents D divided by the number of documents containing the word "w". TF-IDF calculates weights for each word. To determine the TF-IDF value of each sentence, the TF-IDF weights of its constituent words need to be summed. Thus, the weight of a sentence is the sum of the TF-IDF scores of the words within that sentence, as depicted in formula (4).

$$tf - idf(S) = \sum tf - idf(w) \tag{4}$$

Evaluation. The Recall-Oriented Understudy for Gisting Evaluation (ROUGE) Metric is utilized to assess the performance of the proposed method. ROUGE metric is a set of techniques for automatically evaluating the quality of a summary by comparing computer-generated automatic summaries with reference summaries produced by experts [28]. In this research, we utilize ROUGE-1, ROUGE-2, and ROUGE-L. ROUGE-1 deals with unigrams; ROUGE-2 involves bigrams (pairs of words); and ROUGE-L evaluates the sequence of words within the summary. Moreover, in the ROUGE evaluation process, precision, recall, and f-measure are also employed. ROUGE-1 Recall and ROUGE-2 Recall quantify how much human-created summaries are integrated into the computer-generated summaries. Precision measures the relevance of the computer-generated summary. Meanwhile, the F-Measure harmoniously combines both recall and precision [29].

3 Result and Discussion

A total of 62 student assignment documents from the course "Islamic Education" were utilized to assess the performance of the proposed method. Preprocessing is necessary before the data is ready for summarization. The preprocessing stage involves removing unused characters such as #, $, @, &, and others. Additionally, documents originating from PDFs or Word files may sometimes result in poor document readability, leading to the appearance of characters like /t and /n. These characters are removed as they lack meaningful content. Another preprocessing step involves eliminating sentences detected to have fewer than three words, such as sentences like "really?" or "why?", and so on.

The preprocessed data will then serve as the input for the tokenizing and sentence embedding processes using IndoBERT. Figure 2 illustrates the outcome of tokenization and sentence embedding using two sentences. The input sentences are divided into tokens during the token embedding stage, with each sentence starting with the

[CLS] token and ending with the [SEP] token to indicate the sentence boundaries. The token length depends on the number of tokens in the longest sentence. As shown in the illustration in Fig. 2, the second sentence has more tokens than the first one, so the token length follows the second sentence. For the first sentence, padding tokens [PAD] are used to fill the empty spaces. This ensures that both sentences have the same length of tokens (10 tokens). These tokens are then translated into their respective IDs, which are unique numerical representations used by IndoBERT. The ID for the "[PAD]" token is typically assigned as zero. This assists the model in recognizing that these tokens are placeholders for padding and should not contribute to the actual computation. The outcome generated by IndoBERT is sentence embeddings with 768 dimensions, representing each sentence as a vector of 768 values.

The resulting sentence embeddings are vector features used as input for k-means clustering. This clustering technique groups together sentences with similar content, ensuring that sentences with related information are clustered together. Given that each cluster comprises similar sentences, it becomes crucial to choose core sentences capable of representing the information within each cluster effectively. The search for sentence candidates for the final summary uses the TF-IDF weighting method, where sentences achieving the highest TF-IDF values are considered the most informative.

Sentence candidates from each cluster are chosen for the final summary. Thus, the number of clusters (k) significantly influences the summarization outcome, as the number of sentences in the summary depends on the number of clusters. To determine the optimal number of clusters, testing is conducted with k values of 3, 5, 7, 9, 11, and 13. The results of the cluster number's impact on the summary can be seen in

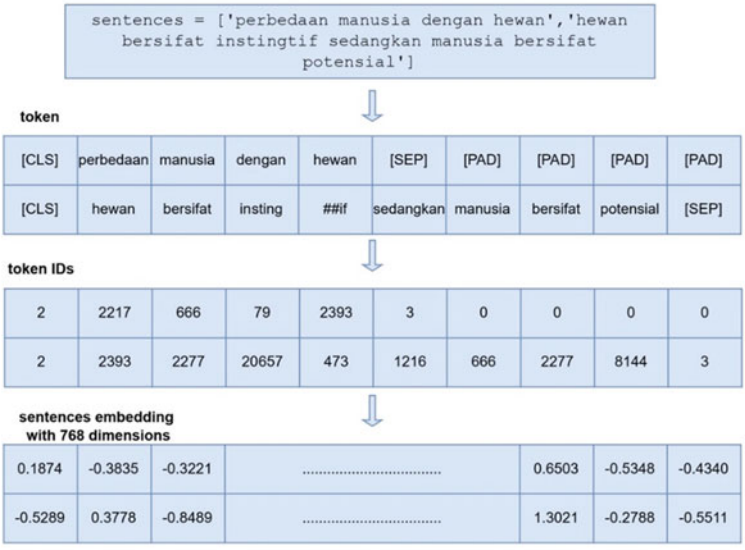


Fig. 2 Illustration of BERT tokenization and output sentence embedding

Table 1. Based on Table 1, the majority of the best ROUGE scores are achieved with k = 13. This implies that the automatic summary result is most relevant to the expert result, with a total of 13 sentences in the summary.

Furthermore, the ROUGE scores achieved through our proposed approach are compared with the ROUGE scores derived from a previous study [17], which employed sentence candidates selected from sentence embeddings closest to the cluster center (centroid). To ensure an objective comparison, the comparison employs the same dataset as utilized in this study. The ROUGE comparison is performed using k = 13 and can be observed in Fig. 3.

Based on the graph in Fig. 3, it can be seen that our proposed method demonstrates slightly improved ROUGE scores. However, when looking at the ROUGE, our proposed TF-IDF weighting based on BERT-k-means reveals no significant gaps when compared to the previous study [17].

The next experiment involves comparing the execution times of the proposed method and the method used in the previous study [17] for summarizing 62 student assignments. As illustrated in Fig. 4, it's clear that the proposed method has a shorter execution time compared to the previous study when summarizing these assignments. This result can be attributed to our method because it has fewer steps than the previous study. In the research conducted by [17], additional steps are introduced, which include employing PCA (Principal Component Analysis) to reduce the number of features in a dataset and Gaussian Mixture, a model for probability distribution. Based on the conducted tests, it can be concluded that the proposed method outperforms the previous study in terms of execution time.

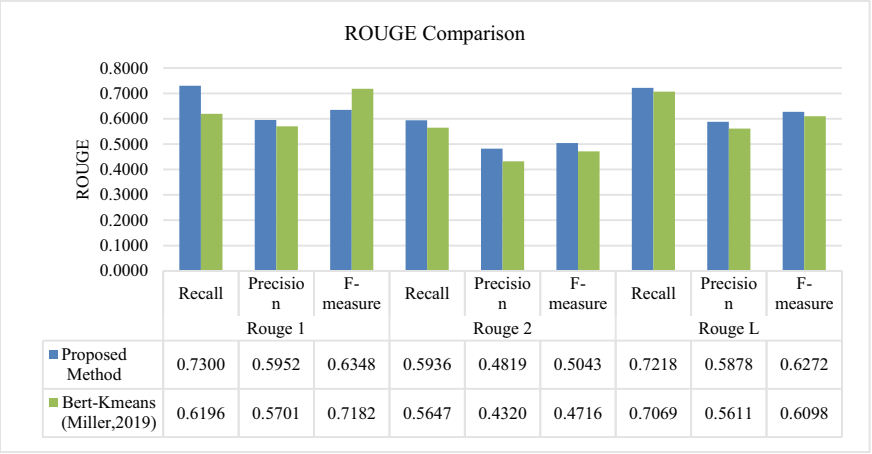
4 Conclusion

In this study, student assignment summarization is performed using BERT and K-means clustering. The study focuses on selecting sentence candidates for the final summary using TF-IDF weighting. BERT is employed to discover contextual information among sentences, and it produces sentence embeddings that represent the sentences. IndoBert, a pre-trained BERT model in Bahasa, is used in this study. The sentence embeddings generated by BERT are then used as input for k-means clustering. Subsequently, TF-IDF weighting is applied to choose the most informative sentences from the clusters. The proposed method achieved the highest ROUGE-1 recall score of 0.73003. While this result cannot be considered optimal, based on the experimental findings, the proposed method can provide slightly better ROUGE scores compared to the previous study. Additionally, we conducted a comparison of execution times for summarizing 62 student assignments, and the results show that our proposed method outperforms the previous study with a shorter execution time.

In the future, there exists potential for further research, particularly in refining the process of selecting sentence candidates from clusters by considering factors such as sentence length, position, or word probability within the title [30].

Table 1 Evaluation result of proposed method

Number of clusters (k)	ROUGE-1			ROUGE-2			ROUGE-L		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
5	0.3824	0.6193	0.4540	0.2380	0.4594	0.2962	0.3720	0.6058	0.4429
7	0.4785	0.6128	0.5186	0.3156	0.4600	0.3544	0.4667	0.5999	0.5068
9	0.5659	0.6147	0.5705	0.4000	0.4775	0.4140	0.5528	0.6017	0.5580
11	0.6641	0.6165	0.6193	0.5073	0.4949	0.4766	0.6522	0.6064	0.6089
13	0.7300	0.5952	0.6348	0.5936	0.4819	0.5043	0.7218	0.5878	0.6272



	Rouge 1			Rouge 2			Rouge L		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
Proposed Method	0.7300	0.5952	0.6348	0.5936	0.4819	0.5043	0.7218	0.5878	0.6272
Bert-Kmeans (Miller,2019)	0.6196	0.5701	0.7182	0.5647	0.4320	0.4716	0.7069	0.5611	0.6098

Fig. 3 Rouge comparison

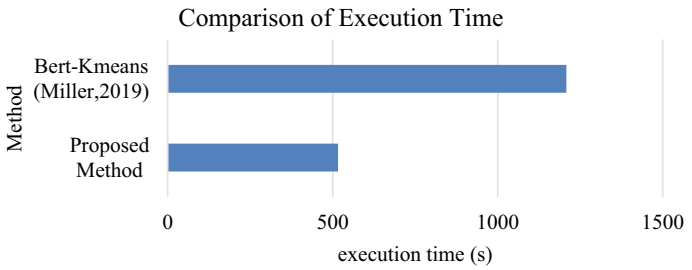


Fig. 4 Execution time comparison

References

1. Kurniawan A, Febrianti AN, Hardianti T et al (2022) Evaluasi Pembelajaran. PT Global Eksekutif Teknologi, Padang
2. Hani'ah M, Kurniawan Y, Rozi IF (2021) livE (onLine--java Exercise) java programming language learning system for lab and online test. Matrix: Jurnal Manajemen Teknologi dan Informatika 11:1–10
3. Setiawati W, Asmira O, Ariyana Y et al (2019) Buku penilaian berorientasi higher order thinking skills
4. Made I, Putra S, Adiwinata Y et al (2021) Extractive text summarization of student essay assignment using sentence weight features and fuzzy C-means. Int J Artif Intell Res 5:13–24. https://doi.org/10.29099/IJAIR.V5I1.187
5. Srivastava R, Singh P, Rana KPS, Kumar V (2022) A topic modeled unsupervised approach to single document extractive text summarization. Knowl Based Syst 246:108636. https://doi.org/10.1016/J.KNOSYS.2022.108636
6. Alomari A, Idris N, Sabri AQM, Alsmadi I (2022) Deep reinforcement and transfer learning for abstractive text summarization: a review. Comput Speech Lang 71:101276. https://doi.org/10.1016/J.CSL.2021.101276

7. Gupta H, Patel M (2021) Method of text summarization using LSA and sentence based topic modelling with BERT. In: Proceedings—international conference on artificial intelligence and smart systems (ICAIS). pp 511–517. https://doi.org/10.1109/ICAIS50930.2021.9395976
8. Giarelis N, Mastrokostas C, Karacapilidis N (2023) Abstractive vs. Extractive summarization: an experimental review. Applied Sci 13. https://doi.org/10.3390/app13137620
9. Kumar Y, Kaur K, Kaur S et al (2021) Study of automatic text summarization approaches in different languages. Artif Intell Rev 54:5897–5929. https://doi.org/10.1007/s10462-021-09964-4
10. El-Kassas WS, Salama CR, Rafea AA, Mohamed HK (2021) Automatic text summarization: a comprehensive survey. Expert Syst Appl 165:113679. https://doi.org/10.1016/J.ESWA.2020.113679
11. Yadav AK, Singh A, Dhiman M et al (2022) Extractive text summarization using deep learning approach. Int J Inf Technol 14:2407–2415. https://doi.org/10.1007/s41870-022-00863-7
12. Gupta S, Gupta SK (2019) Abstractive summarization: an overview of the state of the art. Expert Syst Appl 121:49–65. https://doi.org/10.1016/j.eswa.2018.12.011
13. Shaik T, Tao X, Li Y et al (2022) A review of the trends and challenges in adopting natural language processing methods for education feedback analysis. IEEE Access 10:56720–56739. https://doi.org/10.1109/ACCESS.2022.3177752
14. Moradi M, Dorffner G, Samwald M (2020) Deep contextualized embeddings for quantifying the informative content in biomedical text summarization. Comput Methods Programs Biomed 184:105117. https://doi.org/10.1016/J.CMPB.2019.105117
15. Devlin J, Chang MW, Lee K, Toutanova K (2018) BERT: Pre-training of deep bidirectional transformers for language understanding. NAACL HLT 2019. In: Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies. vol 1, pp 4171–4186
16. Vaswani A, Brain G, Shazeer N et al (2017) Attention is all you need. Adv Neural Inf Process Syst 30
17. Miller D (2019) Leveraging BERT for extractive text summarization on lectures. arXiv preprint arXiv:190604165
18. Koto F, Rahimi A, Lau JH, Baldwin T (2020) IndoLEM and IndoBERT: a benchmark dataset and pre-trained language model for Indonesian NLP. COLING 2020. In: Proceedings of the 28th international conference on computational linguistics. pp 757–770. https://doi.org/10.18653/v1/2020.coling-main.66
19. Shetty K, Kallimani JS (2018) Automatic extractive text summarization using K-means clustering. In: International conference on electrical, electronics, communication computer technologies and optimization techniques (ICEECCOT) 2017 2018-January. pp 881–890. https://doi.org/10.1109/ICEECCOT.2017.8284627
20. Lauriola I, Lavelli A, Aiolli F (2022) An introduction to deep learning in natural language processing: models, techniques, and tools. Neurocomputing 470:443–456. https://doi.org/10.1016/J.NEUCOM.2021.05.103
21. Juarto B, Yulianto (2023) Indonesian news classification using IndoBert. Int J Intell Syst Appl Eng 11:454–460
22. Geetha MP, Karthika Renuka D (2021) Improving the performance of aspect based sentiment analysis using fine-tuned BERT base uncased model. Int J Intell Netw 2:64–69. https://doi.org/10.1016/J.IJIN.2021.06.005
23. Perwira A, Dwitama J, Fudholi DH, Hidayat S (2023) Indonesian hate speech detection using bidirectional long short-term memory (Bi-LSTM). Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi) 7:302–309. https://doi.org/10.29207/RESTI.V7I2.4642
24. Hani'ah M, Fatichah C, Purwitasari D (2016) Reduksi Dimensi Fitur Menggunakan Algoritma ALOFT Untuk Pengelompokan Dokumen. JUTI: Jurnal Ilmiah Teknologi Informasi 14:207. https://doi.org/10.12962/J24068535.V14I2.A573
25. Ririd A, Hani'ah M, Putri I (2020) Analisis Pertumbuhan Balita Menggunakan Algoritma Kmeans++ Untuk Mengetahui Resiko Obesitas. SENTIA 2020 12

26. Abu Nada AM, Alajrami E, Al-Saqqa AA, Abu-Naser SS (2020) Arabic text summarization using AraBERT model using extractive text summarization approach
27. Joshi A, Akram SV, Singh D (2022) TF-IDF and key phrase identification based Malayalam document summarization. In: Proceedings of international interdisciplinary humanitarian conference for sustainability (IIHC). pp 1006–1011. https://doi.org/10.1109/IIHC55949.2022.10060011
28. Lin C-Y (2004) Rouge: a package for automatic evaluation of summaries. In: Text summarization branches out. pp 74–81
29. Harinatha SRK, Tasara BT, Qomariyah NN (2021) Evaluating extractive summarization techniques on news articles. In: Proceedings—2021 international seminar on intelligent technology and its application: intelligent systems for the new normal era (ISITIA). pp 88–94. https://doi.org/10.1109/ISITIA52817.2021.9502230
30. Kumar Y, Kaur K, Kaur S (2021) Study of automatic text summarization approaches in different languages. Artif Intell Rev 54:5897–5929. https://doi.org/10.1007/S10462-021-09964-4/TABLES/2

Topic Modeling of Raja Ampat Tourism on TripAdvisor Sites Using Latent Dirichlet Allocation



Dedy Sugiarto, Dimmas Mulya, Syandra Sari, Anung B. Ariwibowo, Is Mardianto, Muhammad Azka Aulia, Fitria Nabilah Putri, Ida Jubaidah, Arfa Maulana, and Alya Shafa Nadia

Abstract In response to the burgeoning interest in tourist destination sentiment analysis, this study focuses on Indonesia's renowned Raja Ampat. Our primary objective is to employ Latent Dirichlet Allocation (LDA), a topic modeling technique, to delve into sentiments expressed in a corpus of 5,227 TripAdvisor documents concerning Raja Ampat. Leveraging the Vader library for data preprocessing, our analysis reveals a dominant trend of positivity, with 80% of sentiments classified as positive, while 15.8% are neutral, and 4.2% are negative. Through systematic experimentation with Bag of Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF) feature representations, two major topics emerge: 'Environmental Tourism' and 'Snorkeling & Services.' Remarkably, the LDA model attains coherence scores of 0.6005 (BoW) and 0.4148 (TF-IDF), underscoring the robustness of our analysis. These findings not only contribute valuable insights into the prevailing sentiments and dominant topics associated with Raja Ampat tourism but also offer practical implications for the industry. The tourism sector in Raja Ampat can leverage these insights to enhance services and experiences, ultimately facilitating sustainable tourism development in this ecologically vital region. This research thus serves as a significant step toward promoting informed decision-making and improving tourist experiences in Raja Ampat and similar destinations, fostering sustainable growth and preservation of natural beauty.

Keywords Sentiment analysis · Topic modelling · Latent Dirichlet Allocation · Raja Ampat · Tourism

D. Sugiarto · S. Sari · A. B. Ariwibowo · I. Jubaidah
Information System Study Program, Trisakti University, Jakarta, Indonesia

D. Mulya (✉) · I. Mardianto · M. A. Aulia · F. N. Putri · A. Maulana · A. S. Nadia
Informatics Study Program, Trisakti University, Jakarta, Indonesia
e-mail: dimmas.mulya@trisakti.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_32

1 Introduction

Raja Ampat is recognized as a well-known tourist destination located in the West Papua province of Indonesia. This place has achieved international prominence due to the mesmerizing charm of its underwater gardens [1]. Economically speaking, Raja Ampat serves as a valuable asset with significant value within the tourism and hospitality industry as an attractive travel spot.

As per the official website of Raja Ampat Regency's central statistical agency, the cumulative tourist arrivals to Raja Ampat in 2022 amounted to 5,725 individuals, encompassing 4,973 foreign visitors and 752 from the domestic sphere within Indonesia. This data showcases a dramatic escalation, surpassing twofold the preceding year's figure of 2,230 visitors in 2021, which consisted of 697 international tourists and 1,533 local travelers. This pattern highlights a substantial upsurge in global interest, especially post the onset of the COVID-19 pandemic [2].

The Ministry of Tourism and Creative Economy has officially indicated that Raja Ampat is categorized as a priority destination for tourism development [3]. As a consequence, tour guides operating in Raja Ampat should possess a comprehensive comprehension of visitors' perspectives to significantly contribute to the progress of this tourist locale. Given the rapid advancement, the viewpoints and assessments of visitors are accessible on a variety of digital platforms that serve as repositories for visitor reviews of their travel experiences. This phenomenon plays a substantial role in shaping the current magnitude of tourist visits; potential visitors frequently conduct surveys based on these online reviews [4]. Among these platforms, TripAdvisor emerges as a prominent example. The management overseeing tourism and travel in Raja Ampat can ascertain the ongoing discussions among visitors and leverage this insight to make informed decisions, for instance, enhancing or expanding amenities that align with these identified subjects.

Various methods can be utilized to develop sentiment detection models that encompass both reviews and their associated topics, as observed in sentiment analysis for Raja Ampat tourist destinations [5, 6]. Nonetheless, the limitation in these studies lies in their discrete examination of sentiment in reviews. An alternative approach exists that delves into scrutinizing sentiment within each topic across Indonesian tourist destinations [7]. Nonetheless, the difficulty in implementing this approach arises due to the need for labeled data. Labeled data in this case refers to review sentences that have already been annotated with sentiment types in those sentences or can be specifically labeling on particular words to explain the cause of the sentiment [7]. However, obtaining labeled data can be challenging when exclusively collecting reviews from platforms like TripAdvisor and can and it can consume a considerable amount of time and cost to perform manual labeling [8, 9]. Furthermore, sentiment types which are too general and do not provide specific guidance on topics can be challenging for the relevant parties to understand when attempting to make improvements to the tourist destination based on the sentiment analysis result [10].

Thus, an approach is warranted that can unveil both the cause of the sentiment (topic of sentiment) and sentiments itself sans the requirement for labeled data. This objective can be realized through the application of sentiment analysis based on topic modeling [11]. This technique capitalizes on the findings derived from topic identification via topic modeling (LDA) and to drive the ensuing sentiment analysis process utilizing the VADER approach (Valence Aware Dictionary for Sentiment Reasoning) use to support the sentiment analysis of each topic [12–15].

Specifically, our contributions are as follows: (1) We propose a novel methodology that combines Latent Dirichlet Allocation (LDA) for topic modeling with the Valence Aware Dictionary for Sentiment Reasoning (VADER) for sentiment analysis, enabling a nuanced examination of visitor sentiments within distinct topics; (2) We analyze a substantial dataset comprising 5,227 TripAdvisor documents, providing a robust assessment of tourist sentiments and topics associated with Raja Ampat. (3) Our findings offer valuable insights for tourism stakeholders, including the local government and tourism agencies in Raja Ampat, aiding in data-driven decision-making to enhance the tourism sector, particularly vital in the post-pandemic era characterized by evolving travel patterns and preferences.

The results of this study will be utilized as additional information for sentiment analysis in tourism and explain the cause of each sentiment without needing to label sentiment of each aspect or topic. Besides that, this research will also provide crucial information for stakeholders in Raja Ampat, such as the local government of Raja Ampat and Tourism Agencies. This information will aid in calculations and facilitate decision-making concerning the development of the tourism sector, an industry significantly impacted by the number of visitors.

2 Related Works

2.1 TripAdvisor

TripAdvisor is an online travel platform that provides a platform for users to research, review, and rate various travel-related services, including accommodations, restaurants, attractions, and other travel-related businesses. It serves as a crowd-sourced review and recommendation platform, enabling travelers to share their experiences, opinions, and ratings about different destinations and services they have utilized [16]. TripAdvisor allows users to browse and search for reviews, photos, and information about hotels, restaurants, flights, vacation rentals, and other travel-related options worldwide. Additionally, businesses in the travel industry can claim their profiles on TripAdvisor to manage their online presence and engage with customer reviews and feedback. The platform plays a significant role in helping travelers make informed decisions and providing businesses with valuable insights into customer sentiments and preferences within the travel and hospitality sector [17].

2.2 Web Scrapping

Web Scrapping is the automated process of using software to extract specific data from websites. It involves sending requests to web servers, downloading webpage content, parsing the HTML code to find relevant information, and then storing or analyzing the collected data. This technique is used to gather both structured and unstructured data from websites, making it suitable for analysis, research, or integration into other applications. By navigating webpage structures and identifying elements, web scraping eliminates the need for manual data collection and allows for efficient automation of data extraction [18].

2.3 Sentiment Analysis

Sentiment analysis is a computational method in the field of natural language processing and text analysis. It focuses on systematically recognizing, extracting, and categorizing subjective information and emotional expressions found in text. The main aim is to determine the sentiment, whether positive, negative, or neutral, conveyed by a given piece of text [19]. Sentiment analysis has diverse applications, including monitoring social media, evaluating brand perception, summarizing product reviews, and gauging market sentiment. By automating the interpretation of human emotions and viewpoints, sentiment analysis contributes to understanding public sentiment, making well-informed choices, and assessing how products, services, and events are received [20]. There are many approaches in sentiment analysis, such as aspect-based sentiment analysis and lexicon-based sentiment analysis. In this research will focus on usage of lexicon-based sentiment analysis [21].

2.4 VADER

VADER (Valence Aware Dictionary for Sentiment Reasoning) is a sentiment analysis tool designed to evaluate the sentiment of natural language text. Created for sentiment analysis purposes, VADER excels at detecting complex sentiment expressions in text, encompassing both positive and negative sentiments, while also considering the strength of these sentiments. VADER finds widespread utility in areas like social media analysis and evaluating customer feedback, where swift sentiment assessment of text is crucial. Its proficiency in handling informal language and informal expressions makes it highly effective for capturing sentiments in various contexts [12, 22].

2.5 Topic Modeling

Topic modeling is a computational approach in the field of natural language processing and machine learning. It's used to uncover hidden thematic patterns within a collection of text documents [23]. The aim is to identify key topics or themes that contribute to the content of the documents, even without explicit labels [24]. This technique involves using probabilistic models or matrix factorization to analyze word patterns across documents. Through this process, topic modeling helps categorize documents into different topics and provides insights into the main subjects discussed in the text. Extracted topics are represented as word distributions, and this method is widely used for tasks like document grouping, content summarization, and gaining understanding from large text dataset [25].

2.6 Latent Dirichlet Allocation (LDA)

Latent Dirichlet Allocation (LDA) can be described as a probabilistic model extensively used in the realms of natural language processing and machine learning. Its main purpose is to uncover hidden topics within a set of documents and associate words with these topics based on statistical patterns. The model assumes that documents consist of a mixture of topics, and each topic is composed of a mixture of words [13]. LDA aims to reverse this process to reveal the latent topics present in the dataset and the distribution of words connected to each topic. Through methods like variational inference, LDA identifies these topics and their corresponding word distributions, thus enabling the grouping of documents based on their topic makeup. LDA finds application in diverse areas, including text classification [26], content recommendation [27], and the exploration of thematic structures within textual data [28]. It forms a fundamental component of topic modeling, offering insights into the latent topics and themes embedded in extensive collections of text [29].

3 Research Methodology

There are 4 phases carried out in this research. The first step is the data collection phase, followed by pre-processing data phase. Next step is to label the sentiment of each data that had been pre-processed in sentiment analysis phase. The last step is creating the topic modelling model with LDA using the pre-processed data in topic modelling phase. The employed stages are shown in Fig. 1.

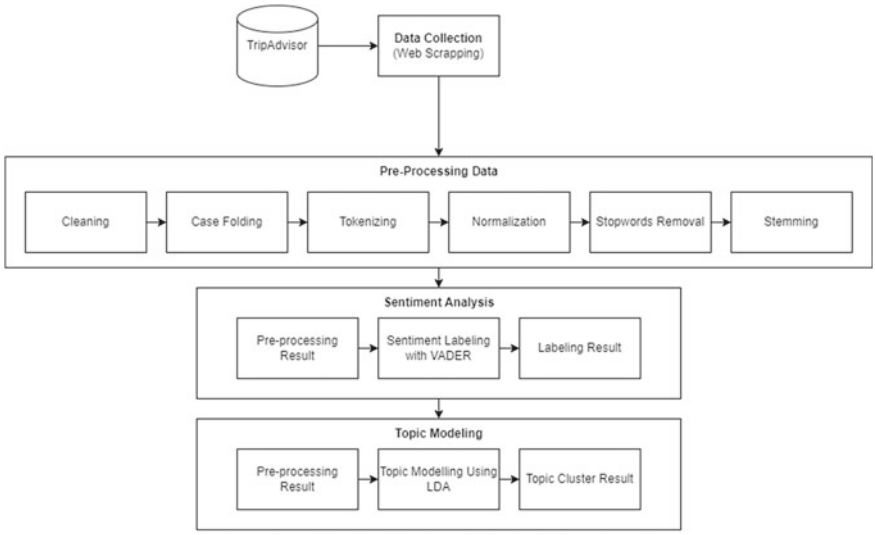


Fig. 1 Research methodology

3.1 Data Collection

The data collection process is carried out using the Web Scraping Technique. The travel review platform targeted for data collection is TripAdvisor, with the focus on Raja Ampat destination. The link which used is tripadvisor.co.id. The gathered information encompasses travel reviews by visitors to Raja Ampat, including details such as review date, rating, title, and review content.

In this research, to execute the scraping process or commonly known as crawling using libraries for data collection, the utilized libraries included *selenium*, *time*, and *csv*. The purpose of the *selenium* library is to fetch data from the TripAdvisor site, the *time* library is employed to introduce pauses so that the program can momentarily halt when the browser is loading a page on the TripAdvisor site, while the *csv* library is valuable for storing the scraping results in csv format. The flow of crawling process is explained in Fig. 2.

In the flow of crawling process, the key to crawling process is in the URL. When the URL page that will be scrapped is already determined, in TripAdvisor there's a keyword of *or10*. That keyword has a function to show the page according to the number of reviews. If the reviews are more than one page, then the keywords will show accordingly with or10 (shown page one), or20 (shown page two), and so on. The crawling process gave result of 5.227 reviews on tourist destination Raja Ampat.

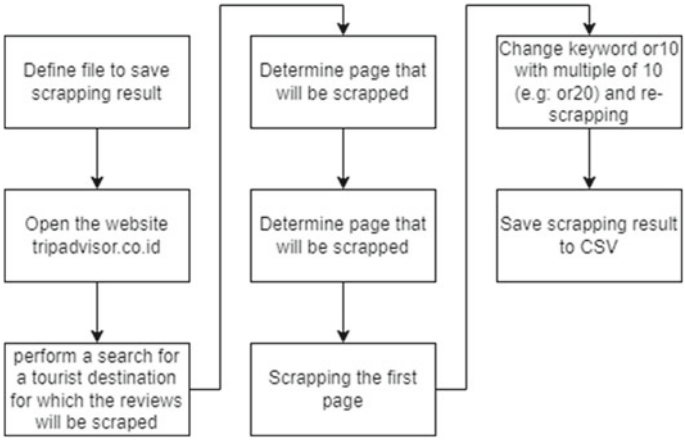


Fig. 2 The flow of crawling process

3.2 Pre-processing Data

Prior to undergoing subsequent processing steps, the data obtained via web scraping will be subjected to preliminary preprocessing to establish appropriate organization and structure. This preliminary processing is essential to arrange the initial data, which exists in varying formats, in a manner suitable for utilization with the classification approach. The predetermined preprocessing procedure is delineated as follows:

- **Cleaning** to remove duplicate data, along with removing punctuation marks, numbers, URLs, and words that have no semantic value.
- **Case folding** to change all the characters in the data to lowercase.
- **Tokenizing** to break sentences into word units or tokens. This is done for all characters and not just words.
- **Normalization** to change non-standard words such as slang words into new words.
- **Stopwords Removal** to remove stopwords, such as conjunctions. This is because stopwords have no semantic value in topic modeling.
- **Stemming** to change words with prefix and suffix to root word.

3.3 Sentiment Analysis

After completing data cleansing the next step involves performing sentiment analysis using VADER. This sentiment analysis process employs the VADER technique, which categorizes the outcomes of each subject into positive, neutral, and negative sentiments. The presentation of these categorizations is visualized using a pie chart to compare the proportions of positive, neutral, and negative sentiments. Additionally,

the WordCloud is utilized to observe the frequently occurring words corresponding to the generated sentiments.

3.3.1 Topic Modeling

Next step is topic modeling using LDA, aiming to establish a model centered around different topics. The outputs from this model creation are then utilized to recognize the prominent subjects present within the document. Following this, the phase of topic validation is executed to guarantee that the model crafted attains the highest probable value. The evaluation hinges on Topic Coherence; achieving a substantial coherence score signifies the efficacy of the resultant model.

4 Result

4.1 Sentiment Analysis

Sentiment analysis was subsequently conducted to observe whether the reviews leaned towards being positive, negative, or neutral. In this sentiment analysis, the VADER tool is used, which is a sentiment analysis algorithm developed to predict the sentiment of text in natural language, such as positive, negative, or neutral.

In this approach, VADER combines various components like a lexicon with its source language, namely *en* (English), word patterns, grammatical processing, and a sentiment scale to analyze text and provide a sentiment score to each given text. This score is then employed to determine whether the text possesses a positive, negative, or neutral sentiment, as well as to measure the intensity level of the sentiment. In this study, the sentiment score ranges from -1 to $+1$, where a negative number signifies negative sentiment, a positive number signifies positive sentiment, and zero signifies neutral sentiment. The result of each sentiment is predominantly positive sentiments (80%), followed by neutral (15.8%) and negative (4.2%) sentiments (Fig. 3a). The visualizations of positive, negative, and neutral WordCloud are presented respectively in Fig. 3b, Fig. 3c, and Fig. 3d. The utilization of WordCloud is exhibited through font size, serving as an indicator of word occurrence frequency. A larger font signifies a more frequent occurrence. In Fig. 3b, the WordCloud that embodies positive sentiment is portrayed. Prominent terms encompass “selam”, “bagus”, “milik”, “terbaik”, “senang”, and “resor”. This signifies the recognition of Raja Ampat’s tourism, particularly the well-regarded diving activities and quality resorts. Figure 3c showcases the WordCloud representing negative sentiment, featuring predominant terms like “air”, “salah”, “selam”, “buruk”, “orang”, “pandang”, and “kotor”. Initially, the keywords in both negative and positive WordClouds might seem almost identical. However, a differentiation lies in the occurrence of terms like “kerusakan” and “kotor”. This suggests that negative sentiment stems from the polluted environment and damage

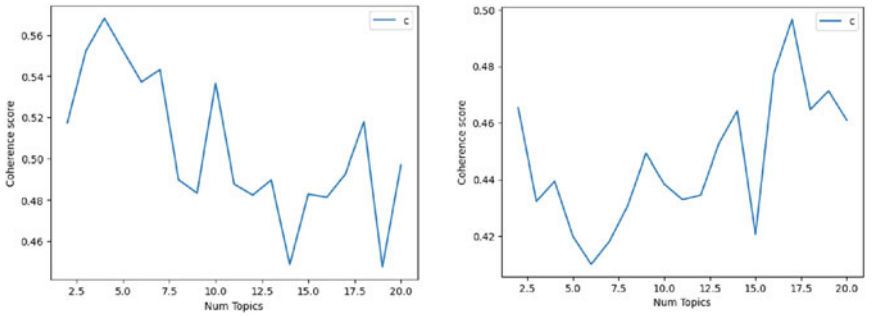


Fig. 3 **a** Distribution of each sentiment labels, **b** WordCloud of positive sentiment, **c** WordCloud of negative sentiment, **d** WordCloud of neutral sentiment

to the used ships or boats. Figure 3d presents neutral sentiment, featuring dominant terms like "pelabuhan", "kota waisai", "makan", "lokasi", "kota", and so forth. WordCloud for neutral sentiment doesn't overtly showcase word dominance due to the relatively uniform font sizes.

4.2 Topic Modeling

In this study, topic modeling employs LDA, where the LDA model incorporates several parameters that need understanding, such as lambda, passes, number of topics, iterations, and feature representations (Bag of Words or TF-IDF). Wherein, lambda represents the optimal topic coherence that has been fine-tuned through manual interpretation. Furthermore, passes determine how many times the entire corpus is



(a) (b)

Fig. 4 **a** Coherence score with bag of words, **b** Coherence score with TF-IDF

traversed during training, while the number of topics is used to confine the analyzed topics, and "iterations" refer to the count of iteration updates performed on each document. In this research, each tuning parameter employs relevance metric = 0.5, alpha = symmetric, passes = 20, number of topics = 2, and iterations = 1000. The numbers of topics are determined from the best coherence score after the experiment process. The highest value of coherence score is the model with four topics. The result of the experiment is shown in Fig. 4a (Bag of Words) and Fig. 4b (TF-IDF) for each type of feature representation.

The highest coherence score from model with Bag of Word feature representation is 0.568146 with four topics. Meanwhile, the coherence score of the model with TF-IDF feature representation is taken only from number of topics under five topics. Even if the coherence score is better at more than 16 topics, it will backfire at analysis complexity for being too many topics that can be comprehended. Within range of one to five topics, the coherence score of the best model with TF-IDF feature representation is 0.465365 with two topics. The best model is built with feature representation type Bag of Words with number of topic equals four. The best model will be used in analysis of all the rest of the sections. The result of the best topic modeling is illustrated in Fig. 5.

During the analysis of the outcomes of topic modeling, it can be discerned that there are four main topics that are being widely discussed in relation to Raja Ampat tourism. These two topics are described in Table 1.

Derived from the four groups of topic keywords, it can be succinctly put that the ongoing prevalent talks center on **the sight of nature** (topic 1) indicated by the terms 'pandang' and 'alam', **the bay** (topic 2) by the terms 'pelabuhan' dan 'bagus', **the food** (topic 3) by the tems 'makanan' and 'enak', and **the snorkeling experiences** described by the words 'selam' and 'pengalaman'.

There are similarities in topic keywords between one topic another because basically Raja Ampat tourist destinations focus on waters and underwater parks which located in Waisai and the main attraction of Raja Ampat is snorkeling. For example,

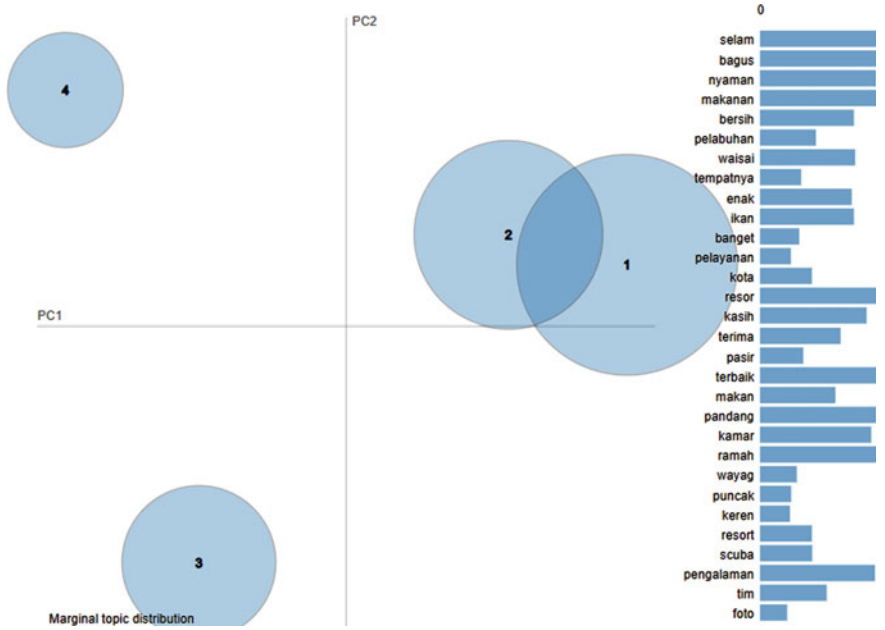


Fig. 5 Topic modeling result with two topics. On the right-hand side of the figure describe the list of words which frequently used in review data in descending order

Table 1 Topic modeling word result with top 15 words

No	Topic 1	Topic 2	Topic 3	Topic 4
1	waisai	bagus	bagus	selam
2	pandang	nyaman	makanan	terbaik
3	air	pelabuhan	enak	kasih
4	salah	bersih	ikan	pengalaman
5	pasir	tempatnya	resor	milik
6	wayang	waisai	ramah	senang
7	alam	banget	kamart	terima
8	surga	pelayanan	selam	orang
9	puncak	kota	makan	tim
10	keindahan	keren	karang	membantu
11	foto	lumayan	air	jalan
12	bagus	cocok	snorkeling	tinggal
13	jalan	pandang	staf	staf
14	spot	mantap	milik	rekomendasi
15	wisata	resort	pandang	penyelam

the word ‘air’ found in topic 1 and topic 3, the word ‘waisai’ found in topic 1 and topic 2, the word ‘selam’ found in topic 3 and topic 4, and so on.

5 Discussion

The establishment of topic modeling and sentiment analysis has been achieved through the LDA method and VADER approach. This technique proves beneficial in grasping sentiment on a per-topic basis from reviews, as illustrated in the context of Raja Ampat tourism. The LDA-based topic modeling results showcase those discussions in reviews prominently involve snorkeling and the sight or the environment of the bay. The positive sentiment contains the sight of nature, the bay, the food, and the snorkeling experiences topics keywords indicates an overall positive response, but at the sight of nature topic keywords appear in negative sentiment seemingly suggests relatively less favorable sentiment. The keyword of the negative sentiment is the word ‘salah’, but after deeper analysis, the word ‘salah’ taken from union words ‘salah satu’, that means *one of* and supposedly not has any sentiment meaning in that word. Hence, the word ‘salah’ has no negative sentiment and that is make all the topics have positive sentiment entirely. A sentiment analysis employing VADER has been applied to discern the sentiment within each review. Crafting WordClouds from positive, negative, and neutral review clusters has enabled the extraction of keywords reflecting both positive, negative, and neutral sentiments. The positive sentiment WordCloud validates the harmony of keywords from each of four topics, affirming the positivity sourced from the thematic basis. In contrast, the negative sentiment WordCloud sheds light on the underlying aspects of the unsatisfied review, however the number of negative sentiments is the most less than other sentiment, so we can consider it as an outlier review about Raja Ampat tourist destination.

The results of the topic modeling and sentiment analysis demonstrate a promising avenue for capturing sentiment within a collection of tourism reviews without the obligation to individually label each sentence’s sentiment. Nevertheless, the study does encounter limitations linked to its dependency on the English-centric VADER library, necessitating the translation from Indonesian. Consequently, forthcoming research endeavors will delve deeper into the available methodologies to alleviate the language-related constraints.

This research offers several advantages compared to sentiment analysis studies in tourist destinations that employ a common approach, which predicts sentiment on a per-sentence basis. This study provides insights based on sentiment generated by topic modeling and sentiment analysis through VADER. In this context, using sentiment data per sentence can be better utilized per topic compared to only predicting general sentiment based on sentences. Each topic’s sentiment can provide more detailed insights and feedback, allowing stakeholders and tourism managers to pinpoint areas for improvement or enhancement in greater detail.

However, the limitation of this study is its inability to provide more detailed sentiment values compared to aspect-based sentiment analysis. This is because this

research can only provide sentiment scores per topic in a general sense, without breaking down sentiment scores into percentages. Additionally, this method still relies on individual or group interpretations, which may introduce bias when applied to specific topics. Nevertheless, at least in this study, we succeeded in pinpointing sentiment within specific tourism topics. This research can be applied to harness sentiment data related to tourism or other topics, with sentiment labels at the sentence level. Consequently, the provided sentiment is not just a general sentiment for a target, but rather relates to topics associated with the sentiment generated by the data.

6 Conclusion

In response to the burgeoning interest in tourist destination sentiment analysis, the research on Raja Ampat tourism's topic modelling-based sentiment analysis was successfully executed using the LDA and VADER techniques. The gathered review data from TripAdvisor underwent preprocessing stages and led to the creation of a topic modeling model aimed at identifying current hot topics and how their sentiment. Sentiment values were assessed via VADER for the reviews. The result of the sentiment by VADER was mostly positive with rate value 80.0% and followed by negative sentiment 4.2% and neutral sentiment 15.8%. After the sentiment analysis step, topics of the review emerged using LDA method with the best model give result of four topics, namely the sight of nature, the bay, the food, and the snorkeling experiences. Sentiment analysis outcomes for all the topics demonstrated an overall positive sentiment, while the negative sentiment seems become an outlier in this review dataset and it has been proven by all the keywords of each topic and deeper analysis of each keyword. Thus, the development of Raja Ampat tourism is advised to maintain the sustainability of the place and improving facilities that become the main attraction to the tourists. Other facilities have garnered positive evaluations and should be preserved to uphold the positive sentiment from visitors.

Future research endeavors will focus on constructing sentiment analysis models that can provide more refined sentiment scores, whether by utilizing pre-trained models or encoders to enhance the accuracy of sentiment results, surpassing the capabilities of VADER alone. Additionally, the experimental design for topic modeling development will be refined to produce smoother and more easily interpretable outcomes. By addressing these aspects, the research aims to deliver superior enhancements and deeper insights.

References

1. McLeod E, Szuster B, Salm R (2009) Sasi and marine conservation in Raja Ampat, Indonesia. Coast Manag 37:656–676. https://doi.org/10.1080/08920750903244143

2. Number of visitors from foreign countries to Raja Ampat in Indonesia from 2013 to 2022. https://www.statista.com/statistics/1002083/international-visitor-numbers-raja-ampat-indonesia/. Accessed 15 Aug 2023
3. Challenges in the development of tourism destinations in the regions
4. Manangiuli L, Sinulingga S, Sibarani R (2019) The influence of tourism product components and promotions on tourist destination image and the impact on revisit intentions to Lake Toba Area, North Sumatra. Saudi J Bus Manag Stud 472–480. https://doi.org/10.36348/sjbms.2019.v04i05.011
5. Laturiuw AK, Singgalen YA (2023) Sentiment analysis of Raja Ampat tourism destination using CRISP-DM: SVM, NBC, DT, and k-NN algorithm. J Inf Syst Inform 5:518–535. https://doi.org/10.51519/journalisi.v5i2.490
6. Nurcahyawati V, Mustaffa Z (2021) Feature selection based on particle swarm optimization algorithm for sentiment analysis classification. In: 2021 International conference on intelligent technology, system and service for internet of everything (ITSS-IoE). IEEE, pp 1–7. https://doi.org/10.1109/ITSS-IoE53029.2021.9615311
7. Huda C, Heryadi Y, Lukas Budiharto W (2022) Aspect-based sentiment analysis in tourism industry for tourism recommender system. In: 2022 5th International seminar on research of information technology and intelligent systems (ISRITI). IEEE, pp 407–412. https://doi.org/10.1109/ISRITI56927.2022.10053083
8. Abonizio HQ, Paraiso EC, Barbon S (2022) Toward text data augmentation for sentiment analysis. IEEE Trans Artif Intell 3:657–668. https://doi.org/10.1109/TAI.2021.3114390
9. Hussein DME-DM (2018) A survey on sentiment analysis challenges. J King Saud University—Engineering Sci 30:330–338. https://doi.org/10.1016/j.jksues.2016.04.002
10. Nasreen Taj MB, Girisha GS (2021) Insights of strength and weakness of evolving methodologies of sentiment analysis. Glob Transit Proc 2:157–162. https://doi.org/10.1016/j.gltp.2021.08.059
11. Ali T, Omar B, Soulaimane K (2022) Analyzing tourism reviews using an LDA topic-based sentiment analysis approach. MethodsX 9:101894. https://doi.org/10.1016/j.mex.2022.101894
12. Hutto C, Gilbert E (2014) VADER: a parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the international AAAI conference on web and social media vol 8, pp 216–225. https://doi.org/10.1609/icwsm.v8i1.14550
13. Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet Allocation. J Mach Learn Res 3:993–1022
14. Yin H, Song X, Yang S, Li J (2022) Sentiment analysis and topic modeling for COVID-19 vaccine discussions. World Wide Web 25:1067–1083. https://doi.org/10.1007/s11280-022-01029-y
15. Çilgin C, Baş M, Bilgehan H, Ünal C (2022) Twitter sentiment analysis during Covid-19 outbreak with VADER. AJIT-e: Acad J Inf Technol 13:72–89. https://doi.org/10.5824/ajite.2022.02.001.x
16. Yoo K-H, Sigala M, Gretzel U (2016) Exploring TripAdvisor. Presented at the https://doi.org/10.1007/978-3-642-54089-9_17
17. Sumarsono D, Sudardi B, Abdullah W (2019) The influence of TripAdvisor application usage towards hotel occupancy rate in Solo. J Phys Conf Ser 1175:012248. https://doi.org/10.1088/1742-6596/1175/1/012248
18. Zhao B (2017) Web scraping. In: Encyclopedia of big data. Springer International Publishing, Cham, pp 1–3. https://doi.org/10.1007/978-3-319-32001-4_483-1
19. Fang X, Zhan J (2015) Sentiment analysis using product review data. J Big Data 2:5. https://doi.org/10.1186/s40537-015-0015-2
20. Wankhade M, Rao ACS, Kulkarni C (2022) A survey on sentiment analysis methods, applications, and challenges. Artif Intell Rev 55:5731–5780. https://doi.org/10.1007/s10462-022-10144-1
21. Firdaus R, Asror I, Herdiani A (2021) Lexicon-based sentiment analysis of Indonesian language student feedback evaluation. Indones J Comput 6
22. Jain R, Kumar A, Nayyar A, Dewan K, Garg R, Raman S, Ganguly S (2023) Explaining sentiment analysis results on social media texts through visualization. Multimed Tools Appl 82:22613–22629. https://doi.org/10.1007/s11042-023-14432-y

23. Blei DM (2012) Probabilistic topic models. Commun ACM 55:77–84. https://doi.org/10.1145/2133806.2133826
24. Liu L, Tang L, Dong W, Yao S, Zhou W (2016) An overview of topic modeling and its current applications in bioinformatics. Springerplus 5:1608. https://doi.org/10.1186/s40064-016-3252-8
25. Churchill R, Singh L (2022) The evolution of topic modeling. ACM Comput Surv 54:1–35. https://doi.org/10.1145/3507900
26. Liu Y, Jin Z (2015) A text classification model constructed by Latent Dirichlet Allocation and Deep Learning. In: Proceedings of the 4th international conference on mechatronics, materials, chemistry and computer engineering 2015. Atlantis Press, Paris, France. https://doi.org/10.2991/icmmcce-15.2015.482
27. Bagul DV, Barve S (2021) A novel content-based recommendation approach based on LDA topic modeling for literature recommendation. In: 2021 6th International conference on inventive computation technologies (ICICT). IEEE, pp 954–961. https://doi.org/10.1109/ICICT50816.2021.9358561
28. Forrer R (2015) The fundamentals of uncovering thematic content in text-based data using LDA. Social Science Research Network
29. Jelodar H, Wang Y, Yuan C, Feng X, Jiang X, Li Y, Zhao L (2019) Latent Dirichlet Allocation (LDA) and topic modeling: models, applications, a survey. Multimed Tools Appl 78:15169–15211. https://doi.org/10.1007/s11042-018-6894-4

Bluetooth and Microcontroller Enabled Wireless Exposure Switch Development for X-ray Mobile Unit to Improve Radiation Protection



Akhmad Haris Sulistiyadi, Dwi Rochmayanti, and Ardi Soesilo Wibowo

Abstract The possibility of radiographers being exposed to harmful scattered radiation in X-ray examinations with X-ray mobile unit is higher because the locations are not possibly designed with protective shielding. The longer distance of the exposure position can minimize the exposure to the radiographer, but is limited to the length of the wire. This study aims to develop a wireless exposure switch with Bluetooth and microcontroller and to verify its performance for safer X-ray mobile examinations. Wireless exposure switch was developed based on Bluetooth version 4 and microcontroller Arduino Nano. Performance tests were performed in the laboratory, consisting of functional test to each component, connection test from 1–15 m, response time from the transmitter to the receiver, and its ability to produce X-ray image. Data were analyzed descriptively to find out the performance. The developed wireless exposure switch can trigger an exposure up to 15 m, 6 times longer than the conventional wire exposure switch. The average response times are 0.061 s for electrons forming and 0.059 s for X-ray forming. It is compatible with an installed X-ray mobile system to produce X-ray images. The Bluetooth and microcontroller-enabled exposure switch are functionally tested and applicable to the X-ray mobile unit. It is prospective to be applied for minimizing radiation exposure for radiographers.

Keywords X-ray exposure switch · Bluetooth · Microcontroller

1 Introduction

Radiation exposure leads to harmful effects, both of stochastic effects and deterministic effects [1]. Radiation exposure in radio-diagnostic services is received not only by patients but also radiographers and other personnel. Through quantitative measurement, previous research [2–4] indicate that the scatter radiation in diagnostic X-ray mobile examination must be a concern. The possibility of radiographers being

A. H. Sulistiyadi · D. Rochmayanti (✉) · A. S. Wibowo
Poltekkes Kemenkes Semarang, 50268 Semarang, Central Java, Indonesia
e-mail: dwirochmayanti@poltekkes-smg.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_33

exposed to scatter radiation in radiographic examinations with X-ray mobile unit is relatively high since the facilities such as Intensive Care Unit or Emergency Unit are not specially designed for X-ray examination [2]. The main concern is the increasing number of examinations required for non-transportable patients in environments where many individuals are present and unprotected as well as to verify whether radiation exposures are within the limits established by national law considering the frequency of this type of examination.

The positioning of radiographer during radiographic examination is critical to radiation protection. All nonessential personnel should be removed from the room, or at least be placed outside of a 6-foot radius during x-ray use. Any personnel within 6 feet of an operating medical fluoroscope must wear lead (or lead equivalent) protective aprons [5]. Radiographers must wear lead aprons as shielding to protect radiosensitive organs and carry out exposures from a minimum distance of 2 m or behind a wall that is still reachable by the length of the wire [2, 4, 6–8]. Even though wearing a lead apron, not all radiosensitive organs can be covered. The thyroid and eyes are often not adequately shielded.

Another attempt to minimize radiation exposure is by using long exposure distance [2, 4, 6, 8–10]. The radiation dose rate in X-ray mobile examination decreased significantly if the distance was longer [2]. But, the exposure distance between radiographer and radiation sources on mobile radiographic examinations is limited to the length of the wire [11].

One alternative solution to solve this problem is the application of remote exposure. Remote controlling in X-ray unit is developed by [11] using Bluetooth and microcontroller Arduino. The system can be used to set the exposure parameters and perform exposure, but the system is not practical in the terms of dimension and compatibility to the installed X-ray. It uses a personal computer and electrical system that replace the wiring system inside the X-ray unit. Wireless exposure switch system currently available as the part of "high end" X-ray units, but it is not applicable for another installed X-ray unit.

Combination of Bluetooth and microcontroller for remote controller is not something new. Previous research developed Bluetooth remote control using smartphone [12], home automation [13, 14], telerobotic controller [15], car remote controller [16], remote controlling and remote data acquisition [17], and remote PC controlling [18]. Almost all of the research utilize smartphone for transmitter [12–16] and earlier works using PDA and cellular phone [17, 18]. The previously developed remote controllers are not designed for X-ray mobile unit exposure.

This study aims to develop a wireless exposure switch with Bluetooth and a microcontroller for a mobile radiographic unit and to verify its performance as an effort to improve radiation protection for radiographers.

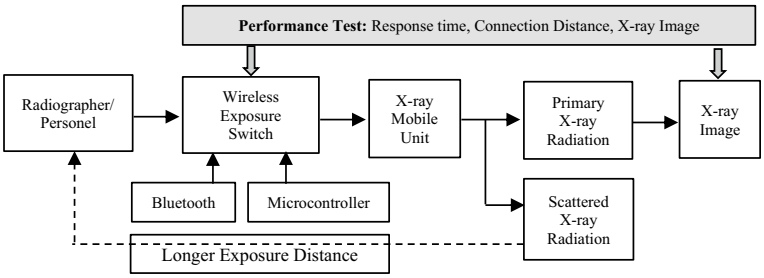


Fig. 1 Theoretical framework

2 Materials and Methods

2.1 Theoretical Framework

This is experimental research with a one-shot case study design. Wireless exposure switch is a special developed exposure switch using Bluetooth and microcontroller that is connected to installed X-ray mobile unit to trigger the X-ray exposure for producing X-ray image. The theoretical framework of the research is shown in Fig. 1.

2.2 Research Procedure

Design and Programming. The system is designed to meet the required features of the radiographic examination using X-ray mobile unit workflow in clinical practice. The electrical circuit is designed to create the subsystems through electrical components. The model of the final shape is carefully designed considering ergonomic and functional aspects. Microcontroller was programmed with the C programming language.

Development. The electrical circuit was developed in the form of a proto-board for trial. Initial testing was performed to verify the design and program. The tested electrical circuit and programmed microcontroller were assembled on a printed circuit board. The 3D-printed case is used to cover the electrical and mechanical parts of the developed wireless exposure switch.

Performance Testing. Performance tests were performed to verify the functions of Bluetooth connectivity, exposure switches, relay systems, power switches, charging systems, indicators lights, and buzzers. Each function are tested until 100 times of constant repeated operation.

Bluetooth connectivity was tested by performing exposure from 1 to 15 m in two conditions: without obstructions and with a 25 cm wall. The successful connection and working system are identified by observation to the LED light and sound indicators (*buzzer*) in the transmitter and receiver module. Test were performed 100 times.

The response times was measured by performing from 1 to 8 m with 5 exposures for each distance. Both of transmitter module and receiver module are connected into a laptop to detect of signal transmitting and signal receiving using free software Arduino Serial Monitor Tool [19]. There are two steps in triggering X-ray exposure using exposure switch: expose preparation (half push) and expose (full push). Both of response times were measured by comparing the transmitting time and receiving time.

The performance in triggering X-ray exposure were confirmed from X-ray image produced with wireless exposure switch that is installed in imaging system (Corsix R32TM Italray X-ray mobile unit and Classic CRTM Carestream CR). A radiographic phantom (Kyoto Kagaku™) was positioned in lateral position to simulate the real examination. Five exposures was performed with exposure factors: 85 kVp, 20 mAs, and 100 cm Source to Image Receptor Distance. Data were analyzed descriptively through the result of the connection and functional test, the response time measurement, and the produced X-ray image.

3 Result

3.1 System Diagram

The developed wireless exposure switch consists of transmitter unit and receiver unit. Transmitter unit consist of charging module, Bluetooth transmitter module, microcontroller, buzzer, and indicator LED. Receiver unit consist of Bluetooth receiver module, microcontroller, buzzer, and indicator LED. The schematic diagram of the system shows in Fig. 2.

3.2 Final Construction

The developed electrical circuit and programmed microcontroller are assembled on a printed circuit board. The 3D-printed case is used to cover the electrical and mechanical parts of the developed wireless exposure switch. The final construction of the wireless exposure switch using Bluetooth and microcontroller is shown in Fig. 3.

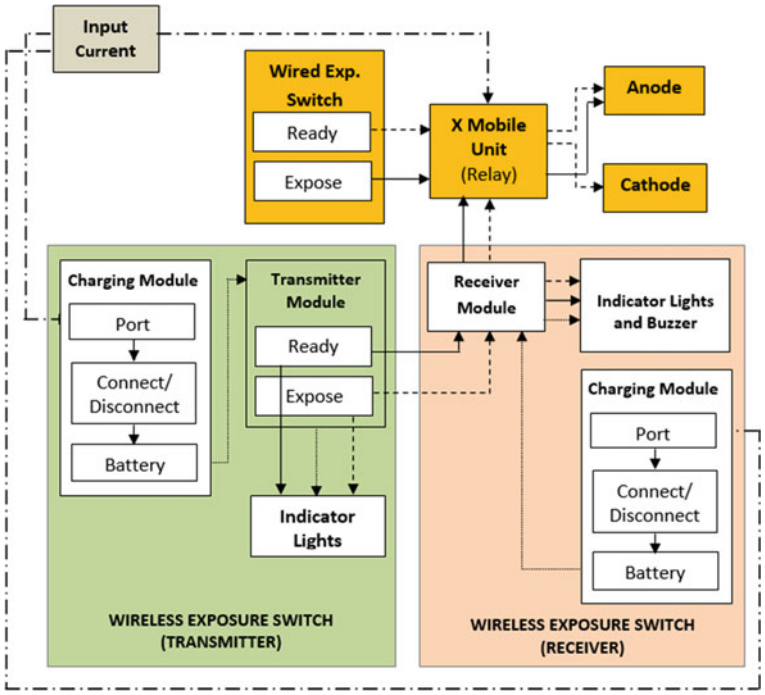
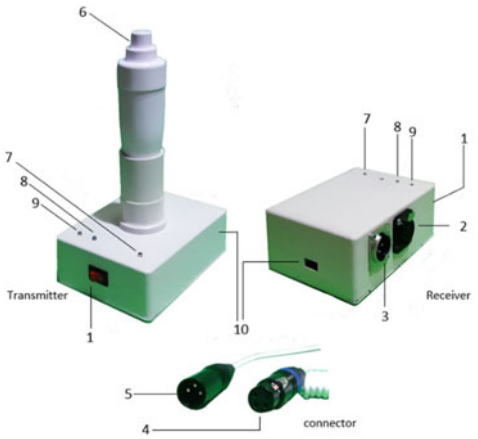


Fig. 2 System diagram of wireless exposure switch for X-ray mobile unit



Annotations:

1. Power button
2. Female socket for exposure switch cable (optional)
3. Male socket for the connector
4. Female socket for the connector, connected to the male socket (3)
5. Male Socket for cable connectors, connected to the socket on the X-ray mobile unit
6. Expose button (ready-expose)
7. Power indicator light (on)
8. Indicator light (expose)
9. Indicator light (ready)
10. Charging port

Fig. 3 Final construction of wireless exposure switch for X-ray mobile unit using Bluetooth and microcontroller

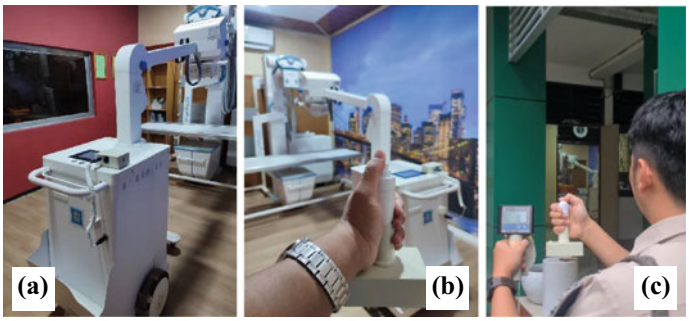


Fig. 4 Operation of the developed Wireless exposure switch. The receiver unit is plugged into the exposure socket in the X-ray mobile unit (**a**). The transmitter unit is hold and pushed as in the conventional one (**b**). The wireless exposure switch can trigger exposure up to 15 m exposure distance (**c**)

Wireless exposure switch is operated by plugging the receiver into the initial socket available in the X-ray mobile unit and pushing the hand switch as shown in Fig. 4.

3.3 Functional Test

Results of the functional test of the Wireless exposure switch for X-ray Mobile Unit using Bluetooth and microcontroller are shown in Table 1.

Table 1 shows that based on 100 times of testing, all functions of the transmitter, receiver, Bluetooth connection, and charging system work properly.

Table 1 Functional test result

Aspect	Testing point	Function
Transmitter	Power button and indicator light turn ON when pressed	Works
	Push button switch triggering exposure and indicator light turn ON and when pressed	Works
Receiver	Power button and indicator light turn ON when pressed	Works
	Buzzer and indicator light: beep and light up when transmitter pressed	Works
Connection	Connected up to 15 m (no barrier)	Works
	Connected up to 3 m (full blockage)	Works
Charging system	Charging when connected to the electrical supply	Works
	Auto-stop charging when the batteries are fully charged	Works

* from 100 testing repetitions

3.4 *Response-Time Test Result*

The response times from 1 to 8 m with 5 in triggering cathode filament heating and triggering X-ray exposure. The results of the response time between the transmitter and receiver are shown in Table 2.

Table 2 shows the response time based on 5 times testing for distance 1 to 8 m. The average response time from 1 to 8 m distance are 0.061 s for filament cathode heating and 0.059 s for X-ray exposure. Standard deviation of the response times in Triggering Electrons Forming and Triggering X-ray forming from 8 variations of distance are ± 0.020 s and ± 0.018 respectively. Figure 5 describes the trend of response time in various exposure distance. The response times are not related to the distance in specific manner.

Table 2 Response test result

Distance (meters)	Response time (second)			
	Triggering filament heating		Triggering X-ray exposure	
	Response*	Std. Dev	Response*	Std. Dev
1	0.053	± 0.017	0.054	± 0.023
2	0.081	± 0.009	0.073	± 0.015
3	0.077	± 0.048	0.076	± 0.068
4	0.047	± 0.000	0.064	± 0.023
5	0.036	± 0.020	0.033	± 0.019
6	0.037	± 0.050	0.031	± 0.022
7	0.083	± 0.050	0.075	± 0.063
8	0.076	± 0.066	0.066	± 0051
Average	0.061	± 0.020	0.059	± 0.018

* average value from 5 times testing

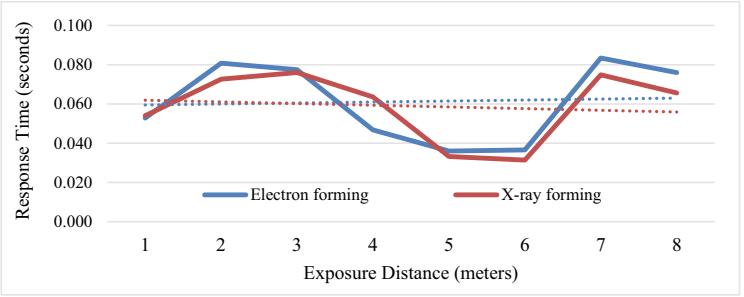


Fig. 5 Response time between transmit time and receive time in various exposure distance



Fig. 6 X-ray image of the lumbar spine produced with wireless exposure switch that installed in the imaging system

3.5 Radiographic Image

The performance of wireless exposure switch in triggering X-ray exposure were confirmed from its ability to produce X-ray image. Figure 6 visualizes a radiographic image of lumbal spine that acquired using wireless exposure switch using installed X-ray mobile unit.

4 Discussion

The principle of exposure to a mobile radiographic unit is basically the same as the stationary unit. The exposure is performed with a switch to trigger the formation of X-rays by the formation of electron clouds on the filament cathode and accelerate toward the anode with a certain tube voltage. The filament cathode heating is done by pressing the exposure button halfway to heat the filament cathode and followed by pressing it fully to release the X-ray from the anode.

Mobile radiographic examinations are usually performed in locations or facilities that are not specifically designed for radiographic examinations, such as in emergency units and intensive care units. The radiographer wears a lead apron and performs the exposure from a position far from the radiation source. The radiation dose rate in X-ray mobile examination decreased significantly if the distance was longer [2, 5]. To minimize radiation exposure radiographers can stand in long exposure distance [6, 7]. However, this method is limited to the length of the exposure switch wire which is around 1–2 m.

Our developed wireless exposure switch allows exposure from long distance to minimize the scattered radiation received by radiographer. It consists of two Bluetooth modules: transmitter and receiver. To trigger the exposure, a relay is used. The operation of the components on is controlled by the Arduino Nano microcontroller.

This component is installed both on the transmitter unit and receiver unit. In the transmitter unit, the microcontroller process instructions in the form of pressure from a push button. These instructions sent to the Bluetooth transmitter module which will transmit instruction to the receiver unit. The Bluetooth module on the receiver unit is connected to the microcontroller to receive instructions and process them to control the relay. The relay is connected to the X-ray generation system on the mobile radiography unit. The Wireless Exposure Switch powered by 99.000 mAh Lithium-Ion battery to supply sufficient power to drive the relay. The smaller Lithium Ion 1200 mAh battery is used in the transmitter. Both batteries are rechargeable. To be fully charged, the transmitter battery takes 23 min and the receiver battery takes 14 min. and can be used up to 60 times exposures. Functional test result shows that all of the Bluetooth connectivity, exposure switches, relay systems, power switches, charging systems, indicators lights, and buzzers work properly to trigger exposure replacing the conventional wire exposure switch.

This wireless exposure switch is enabled with Bluetooth as transmitting media because of its several advantages. According to [20], Bluetooth is a low-cost device. Bluetooth allows data and voice transmission, with the configuration of 1 asynchronous data channel and 3 synchronous voice channels or 1 channel which is a combination of data and voice. The transmission distance is possible up to 10 m and can be increased by increasing the transmission power. The advent of Bluetooth technology has led to the development of various applications, ranging from wireless audio streaming to IoT devices [21–27].

In parallel, Arduino Nano, a compact microcontroller board based on the ATmega328P chip is popular due to its ease of use and versatility. Combining Bluetooth and Arduino Nano offers a powerful platform for building wireless remote-control systems for a wide array of applications. The microcontroller used is the Arduino Nano because of its small size and the instructions it processes are relatively simple. Microcontrollers are widely used for various purposes.

The implementation of Bluetooth and microcontroller for remote controller previously developed in remote control using smartphone [12], home automation [13, 14], telerobotic controller [15], car remote controller [16], remote controlling and remote data acquisition [17], and remote PC controlling [18]. Almost all of the research utilize smartphone for transmitter [12–16] and earlier works using PDA and cellular phone [17, 18]. The previously developed remote controllers are not designed for X-ray mobile unit exposure.

Remote controlling system in X-ray unit are developed by [11] using Bluetooth and microcontroller Arduino. The system can be used to set the exposure parameters and perform exposure, but the system is not practical in the terms of dimension and compatibility to the installed X-ray. It uses a personal computer and electrical system that replaces the wiring system inside the X-ray unit. A wireless exposure switch system currently available as the part of “high end” X-ray units, but it is not applicable to other installed X-ray units.

Unlike the previous system [11], our developed wireless exposure switch is more practical to use. The users only need to plug the receiver unit to the available socket of the conventional wire exposure switch in the X-ray mobile unit. There is no

additional internal electrical modification needed. The conventional wire exposure switch still can be used since the receiver unit is equipped with an additional socket for simultaneous operation. It has higher compatibility than the commercially wireless exposure switch that is only coupled to the specific type and series and sold in unity.

The results of this study indicate that the wireless exposure switch can be connected up to 15 m distance, 7.5 times longer than conventional wire exposure switch. This is very beneficial because the exposure rate is affected by distance. The longer the distance, the smaller the exposure rate according to the best distance square law. The results of research by [2] showed that the exposure rate at a distance of 0.4 m was 0.8 μSv/h and decreased exponentially to 0 μSv/h at a distance of 2.5 m. Ionizing radiation must be a concern because there are stochastic effects that don't have the dose threshold. Small radiation exposure can cause a stochastic effect or no stochastic effect. The same thing happens with exposure to large amounts of radiation [1].

The response time result shows that the delay between transmitting and receiving time is very short, less than one second in the operating distance range. This spontaneous response is very needed to avoid blurred X-ray images due to patient movement to achieve high image quality.

The limitation of our developed wireless exposure switch is only compatible to three-pins socket. For other type of socket, connector in the receiver unit need to be modified and microcontroller need to be reprogramed. It is very possible because the transmitter and receiver unit have a socket for cable-data transfer connection.

The results indicate that our developed wireless exposure switch is functionally tested, responsive, and visible to be applied for improving radiation protection by minimizing the radiation exposure to the radiographers.

5 Conclusion

The wireless exposure switch for X-ray mobile unit developed based on Bluetooth and microcontroller Arduino. All of the functional test is passed. It can trigger exposure from 15 m, 7.5 times longer than conventional wire exposure switch. The scattered radiation exposure to the radiographer can be reduced to the longer exposure distance. The response times are very short, 0.061 s for filament cathode heating and 0.059 s for X-ray exposure, minimizing image blurring due to patient movement during exposure. However, its performance is need to be tested using by comparing the X-ray output and image quality through quantitative test in the further research.

References

1. Clement CH, Stewart FA, Akleyev AV, Hauer-Jensen M, Hendry JH, Kleiman NJ, Macvittie TJ, Aleman BM, Edgar AB, Mabuchi K, Muirhead CR, Shore RE, Wallace WH (2012) ICRP

publication 118: ICRP statement on tissue reactions and early and late effects of radiation in normal tissues and organs Ã¢â‚¬â€ threshold doses for tissue reactions in a radiation protection context. Ann ICRP 41:1–322. https://doi.org/10.1016/j.icrp.2012.02.001
2. Abrantes AFL, Rebelo C, Sousa P, Rodrigues S, Almeida RP, Pinheiro J, Azevedo K, Ribeiro L (2017) Scatter radiation exposure during mobile X-ray examinations. HealthManagement org 17:68–72
3. Kendrick DE, Miller CP, Moorehead PA, Kim AH, Baele HR, Wong VL, Jordan DW, Kashyap VS (2016) Comparative occupational radiation exposure between fixed and mobile imaging systems. J Vasc Surg 63:190–197. https://doi.org/10.1016/j.jvs.2015.08.062
4. de Castro Bienert IR, Ferreira LC, da Silva PA, Miwa DTK, Florêncio CLB, Mota RL, Rinaldi FS, de Andrade PB (2016) Evaluation of scattered radiation and impact of local protective devices in an interventional cardiology laboratory. Rev Bras Cardiol Invasiva (English Ed) 24:38–43. https://doi.org/10.1016/j.rbciev.2017.10.008
5. Mitchell EL, Furey P (2011) Prevention of radiation injury from medical imaging. J Vasc Surg 53:22S-27S. https://doi.org/10.1016/j.jvs.2010.05.139
6. Bontrager KL, Lampignano JP (2014) Textbook of positioning and related anatomy
7. Long B, Rollins J, Smith B (2016) Merrill's pocket guide to radiography e-book
8. Clark KC (2016) Positioning in radiography. Ilford Limited, London
9. Robert Fosbinder DO (2012) Esentials of radiologic science. Lippincott Williams & Wilkins, Philadelphia
10. Bushong SC (2013) Radiologic science fo technologists: physics, biology, and protection. Elsevier Inc., St. Louis, Missouri
11. Wijaya NH, Yudhana A, Robiyansah, Sukwono D (2021) X-ray machine control with wireless based on mA parameters. IOP Conf Ser Mater Sci Eng 1088:012080. https://doi.org/10.1088/1757-899x/1088/1/012080
12. Puthanial M, Rajeshwari S, Raja PCK, Shankar P (2014) Android and bluetooth technology enabledremote control using smart phone. Int J Adv Res Electr Electron Instrum Energy 3:9373–9380
13. Kavitha J, Naveen O, Manoj Kumar PKSR (2018) Bluetooth based home automation using Arduino and android application. Grad Res Eng Technol 6:37–42 https://doi.org/10.47893/gret.2021.1116
14. Sullivan D, Chen W, Pandya A (2017) Design of remote control of home appliances via bluetooth and android smart phones. In: 2017 IEEE international conference on consumer electronics—Taiwan, ICCE-TW 2017. pp 371–372. https://doi.org/10.1109/ICCE-China.2017.7991150
15. Andi A, Zainul I, Julpri A (2018) Design of Telerobotic system using bluetooth with behaviour-based control architecture. IOP Conf Ser Mater Sci Eng 453. https://doi.org/10.1088/1757-899X/453/1/012067
16. Rai N, Rasaily D, Wangchuk TR, Gurung M, Khawas RK (2016) Bluetooth controlled car using Arduino. Int J Eng Trends Technol 33:381–384. https://doi.org/10.14233/ajomc.2022.7-1-sp4.pp135-138
17. Saadat N, Yaacob H, Kalos R, Khatun S, Ali BM, Azmir RS (2006) Bluetooth based wireless remote device controlling and data acquisition. In: Proceedings of advanced international conference on telecommunications international conference on internet web applications and services (AICT/ICIW'06). pp 75. https://doi.org/10.1109/AICT-ICIW.2006.56
18. Alhakim MM, Al-Kittani I, Bakleh A, Swidan M, Zarka N (2006) Bluetooth remote control. In: Proceedings—2006 International conference on information and communication technology: from theory to application (ICTTA 2006). vol 2, pp 2674–2677. https://doi.org/10.1109/ICTTA.2006.1684832
19. Söderby K (2023) Using the serial monitor tool. https://docs.arduino.cc/software/ide-v2/tutorials/ide-v2-serial-monitor. Accessed 08 Nov 2023
20. Sairam KVSSSS, Gunasekaran N, Rama Reddy S (2002) Bluetooth in wireless communication. IEEE Commun Mag 40:90–96. https://doi.org/10.1109/MCOM.2002.1007414

21. Arduino Nano. https://docs.arduino.cc/hardware/nano?queryID=undefined. Accessed 08 Nov 2023
22. Ali HM, Hashim Y, Al-Sakkal GA (2022) Design and implementation of Arduino based robotic arm. Int J Electr Comput Eng 12:1411–1418. https://doi.org/10.11591/ijece.v12i2.pp1411-1418
23. Tounsi M, Qureshi B (2008) A bluetooth-enabled mobile intelligent remote healthcare monitoring system: analysis and design issues. Int J Healthc Technol Manag 9:473–484. https://doi.org/10.1504/IJHTM.2008.020199
24. Chaari Fourati L, Said S (2020) Remote health monitoring systems based on bluetooth low energy (BLE) communication systems. Springer International Publishing. https://doi.org/10.1007/978-3-030-51517-1_4
25. Alshazly H, Hassaballah M, Alshazly HA, Hassaballah M (2016) An embedded system for a bluetooth controlled mobile robot based on the ATmega8535 microcontroller image features description and mathing view project ear recognition view project. Egypt Comput Sci J 40:61
26. Omran MA, Saad WK, Hamza BJ, Al-Baghdadi AF (2021) Designing and manufacturing of home automation monitoring system using internet of things technology. J Phys Conf Ser 1973. https://doi.org/10.1088/1742-6596/1973/1/012081
27. Al-Fuqaha A, Guizani M, Mohammadi M, Aledhari M, Ayyash M (2015) Internet of things: a survey on enabling technologies, protocols, and applications. IEEE Commun Surv Tutorials 17:2347–2376. https://doi.org/10.1109/COMST.2015.2444095

State of the Art of Battery Swap Station Management System in Indonesia: An IoT-Based Prototype



Ridlho Khoirul Fachri, Rashad Abul Khayr, Muhammad Zakiyullah Romdlony, Aam Muharam, and Amin

Abstract This paper discusses a comprehensive review related to Battery Swap Station (BSS) technology in Indonesia and proposes the prototype of an IoT-based battery swap station management system in Indonesia. The spread of BSS in Indonesia can be reviewed in terms of several things that support the development of BSS in Indonesia, such as the type of batteries and technology used on BSS. Some types of batteries can be distinguished according to the brands that exist in Indonesia, including Gogoro, Viar, Smoot, Gesits, and Volta. Battery technology can be reviewed from the connector or plug on the BSS that connects the battery, Battery Management Systems (BMS), and the app to the user. By 2022, there are 369 BSSs scattered across Indonesia, which will continue to be pushed by the Indonesian government for the use of more massive battery-based electric motor vehicles. This leads to the heterogeneity issue of several different battery voltages in the Indonesian market that the suggested prototype can address.

Keywords Internet of things (IoT) · Battery swap station (BSS) · Heterogenity

1 Introduction

Air pollution in the form of carbon dioxide generated by two-wheel motor vehicles and oil-fueled four-wheelers (BBMs) is the largest contributor, accounting for more than 80% of the formation of greenhouse gases that will result in rising global atmospheric temperatures [1–4]. By 2020, the number of motor vehicles in Indonesia will reach 136 million, with 15 million passenger cars, 233 thousand buses, 5 million goods cars, and 115 million motorcycles [5]. The huge number of two-wheeled motor vehicles is one of the biggest contributors to greenhouse gases in Indonesia.

R. K. Fachri (✉) · R. A. Khayr · M. Z. Romdlony
School of Electrical Engineering, Telkom University, Bandung, Indonesia
e-mail: ridlhokhf@student.telkomuniversity.co.id

A. Muharam · Amin
National Research and Innovation Agency (BRIN), Jakarta, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_34

Indonesia has implemented motor vehicle emission regulations to reduce the emissions of the motor vehicle sector by implementing CO_2 reduction in the transport sector, i.e., the use of electric vehicles. Electric vehicles began to flourish again with the issue of Presidential Regulation 55/2019 on battery-based electric vehicles. In addition, various derivative policies have been made by the local government to support the acceleration of the development of electric vehicles. This 55/2019 push has driven the development of the population of electric vehicles to increase rapidly by 2020; some electric vehicles, in particular electric motorcycles, have been made in the country [6].

The infrastructure for electric cars includes charging stations. Indonesia's public electric vehicle charging infrastructure has three parts: the Public Electric Power Supplier Station, which charges electric vehicles, especially electric motorcycles; the Public Electric Vehicle Charging Station, which charges electric cars; and the Battery Swap Station (BSS) [6]. The Ministry of Energy and Mineral Resources Republic of Indonesia released government rule 13/2020 on battery-based electric motor vehicle charging infrastructure in 2020. This government legislation covers electric car charging infrastructure, power rates, and security [6, 7].

The reason of this paper is to obtain the latest status of the development of public electric vehicle charging stations in Indonesia in particular the charging of general electric vehicle type BSS for two-wheeled vehicles. And we also propose a prototype using readings from battery parameters using the Antares platform and the device application and Human Machine Interface (HMI) with display for two types of voltage (heterogeneity) 60 and 72 V. Because of the variation in the voltage (heterogeneity) of the electric battery that is on the Indonesian market and the lack of BSS that can charge two types of battery voltage at once, so we made this prototype with the purpose as a form of one of the contributions to Indonesia. This information and prototype is expected to help the community of electric vehicle users, for research and development in the BSS field in Indonesia.

2 Development of BSS in Indonesia

The battery replacement process on the BSS takes only 5–10 s, so it is quite short. BSS [8–11] is an effort to accelerate the development of electric vehicles. The technical specifications of BSS are still in the planning and preparation phase, which includes two SNI plans: 1. Safety requirements 2. Dimensions of batteries that can be removed and exchanged for electric motor vehicles. Related to battery swap standards, some of the technical aspects discussed are the battery container, nominal voltage (48, 60, and 72 V), battery capacity (14, 20, and 23 Ah), battery connectors, and communication protocols. The Ministry of Energy and Mineral Resources Republic of Indonesia is also in the process of revising the Regulation Number 13 of 2020 in an effort to accelerate the number of BSS spreads through licensing processes, more flexible technical requirements, and other incentives. By 2022, there will be 369 BSS units scattered across Indonesia [7].

3 BSS Technology

Batteries are a very important component for electric vehicles. Like oil on fueled vehicles, fossil fuels are the primary source for vehicles so that they can be used. The batteries used in electric vehicles in Indonesia are specifically for two wheels. There are several types and brands, including Gogoro, Viar, Smoot, Gesits, and Volta. And below in Table 1 is a comparison of specifications for the type of battery reviewed from dimensions (cm) to length (L) × width (W) × height (H), mass (kg), DC voltage (V_{DC}), capacity (Ah and Wh), and mileage (km) in full battery conditions.

In addition to the battery specifications, two-wheel electric motor vehicles require supporting technology to be used on several brands of electric motor batteries circulating in Indonesia. These differences are a determining factor in the needs of society. Some of the technologies available on two-wheeled electric vehicles are connectors or plugs from BSS that connect batteries to BSS, Battery Management Systems (BMS) [17–20], and appliances for consumers [21].

The connector, or plug, is a part to connect between batteries with BSS for the purpose of charging exhausted batteries. Here is a display of the connector or battery plug of several brands of batteries in two-wheel electric motor vehicles described in Table 2. Two-wheel electric motor cars with battery monitoring technology. Battery Management Systems (BMS) are this technology [22, 23] and information BMS may provide: Charge balancing to avoid overload damage by charging all batteries simultaneously; Active balancing shifts energy from stronger to weaker cells to ensure all cells reach the maximum exhaust point at once; Temperature monitoring

Table 1 Battery specifications [12–16]

Name of batteries	Gogoro	Viar	Smoot	Gesits	Volta
Dimensions L×W×H (cm)	18 × 18 × 45	22 × 22 × 50	177 × 177 × 330	194 × 67x113	240 × 150 × 20
Mass (kg)	9.8	11	12	8	11
Voltage (V_{DC})	43	60	64	72	60
Capacity (Ah)	30.3	23	22,5	19,4	23
Capacity (Wh)	1300	1380	1440	1398	1380
Distance to vehicle (Km)	85	50	60	50	60
Time to charge the battery until fully charged (hour)	3–4	4–5	8	3–4	8
Type of battery used	Panasonic 18,650 Lithium ion	Lithium ion 26,650	Lithium ion	Panasonic 21,700 Lithium ion	Lithium ion

to prevent heat damage; Low-voltage cut-off, how to disconnect the battery when any cell exceeds the acceptable minimum voltage, and overuse damage; SoC and battery capacity monitoring. By monitoring voltage and current, cell capacity may be estimated; Battery coordinate position (GPS module); Health of the battery; Power switch protection network (when charging fully) and battery series number information [24, 25].

The information obtained by BMS will be transmitted using Internet of Things (IoT) [22, 26–28] technology and can be forwarded to the app display by connecting via Bluetooth communication. To support the connection of information obtained from BMS to practical consumer applications, BMS is supported by communication protocols [29, 30]. The BMS standard applied is SMBus (System Management Bus), which is used for most portable applications [31], CAN Bus (Controller Area Network), which is an electronic control used to obtain data on vehicles such as motor speed, battery condition, and GPS [32], and LIN Bus (Local Interconnect Network), which is simpler for battery pack use [33, 34].

4 Proposed Prototype

Two types of batteries will be tested. The first type of battery is a battery that has been used by an electric motorcycle, and the second type is a packaged battery adapted to the battery used by the electric motorcycle. Batteries used by motorcycles are equipped with Controller Area Network (CAN) communication lines. The CAN track on the battery is used by motorcycles to monitor the condition of the batteries while the motorcycle is running. The battery condition can also be monitored at the time of charging the battery using the same communication line. When BMS is working, it will broadcast the battery condition through the CAN channel. The broadcast battery condition will be divided into four groups of data that have different identifications. This identification will always be the same for a battery, but the identification of a data group may differ when compared between one battery and the other, even though the two batteries are motorcycle batteries. Despite the differences, battery

Table 2 Form of connector or battery plug [12–16]

Name of batteries	Connector and plug
GOGORO (Put the connector under the battery)	

(continued)

Table 2 (continued)

Name of batteries	Connector and plug
Viar (Put the plug on the battery)	
Smoot (Put the connector under the battery)	
Gesits (Put the plug on the battery)	
Volta (Put the plug on the battery)	

identification is always done in a sequence where the first group contains the general condition of the battery, the second group contains the identifier, and the third group includes the number of battery charging cycles. Figure 1 is a setup of the experiment we did related to the prototype.

Figure 2 shows the readings from the battery that we have made using the Antares database platform with parameters that have been successfully read. A support that is built to carry out one of the primary duties of identifying the position of the battery Gesits exchange station is shown in Fig. 3. This companion application may display the battery state at the battery swap station in addition as displaying the location of the battery exchange station. This program can record information gathered from Internet of Things (IoT) devices.

The data displayed by this app will assist the station owner in monitoring the battery condition. However, the data shown is still too detailed for electric motorcycle



Fig. 1 The experimental setup of the prototype shows that what we did with batteries with different voltage types of 60 and 72 V

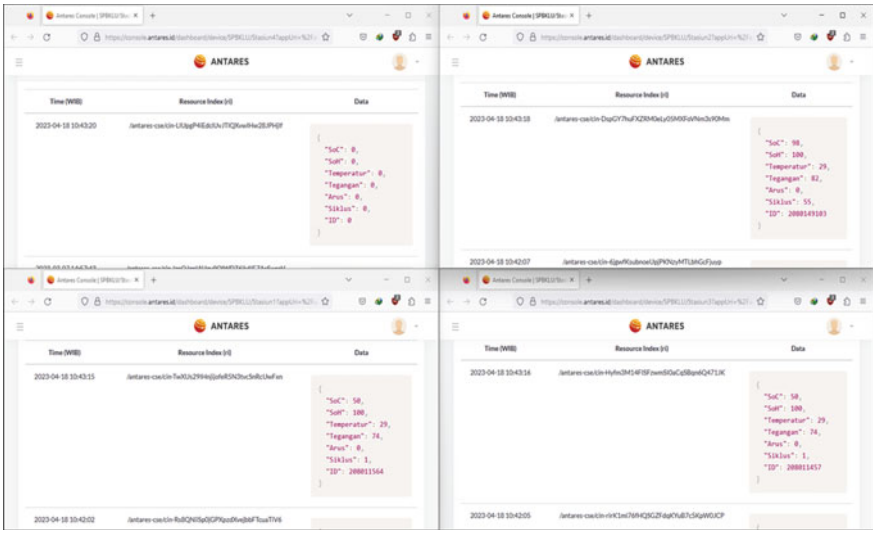


Fig. 2 Readings from the battery using the Antares platform show several parameters, including: state of charge (SoC), state of health (SoH), temperature, current, cycle, and battery identity number (ID)

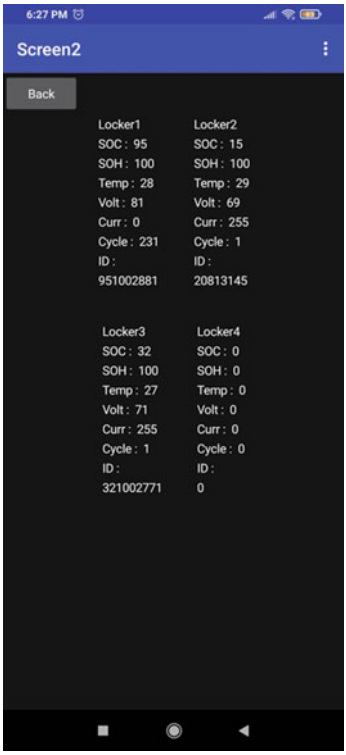


Fig. 3 Menu display on some battery parameters data using the device application created. By showing the success of reading parameters the same as those in the database. The value of the parameters that shows the number 0 is the absence of the battery that is being charged on BSS

users. In addition to data capture, the app is also equipped with communication with BSS for battery conversion transactions. Some of the batteries that have been read in the database can also pass these data parameters to the Human Machine Interface (HMI) that we've created. This HMI can be shown in some of the pictures below, namely from Fig. 4a, b, and c.

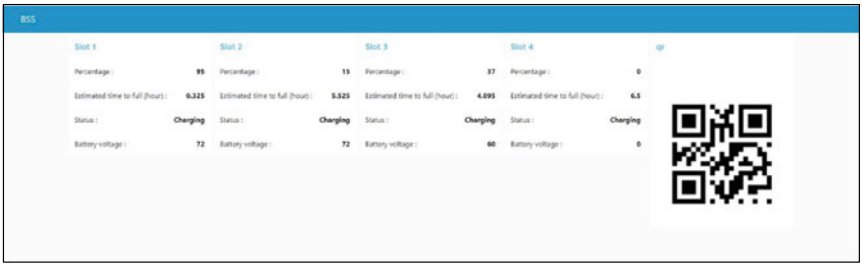
The packaged battery will have a station and HMI that are slightly different from the station for Gesits batteries. On heterogeneity showing HMI with voltage difference can be seen in Fig. 4c and Fig. 5.



(a)



(b)



(c)

Fig. 4 **a** HMI looks from the battery Gesits. This transaction can be done by scanning the QR code displayed on the HMI on the BSS, **b** After the user scans, the HMI will show the removable battery and the available battery slots. Slots looking that can be taken and inserted on HMI, **c** The battery has two types (heterogeneity) of voltage: 72 and 60 V. Both types will be inserted into the same station that will display the type of battery voltage

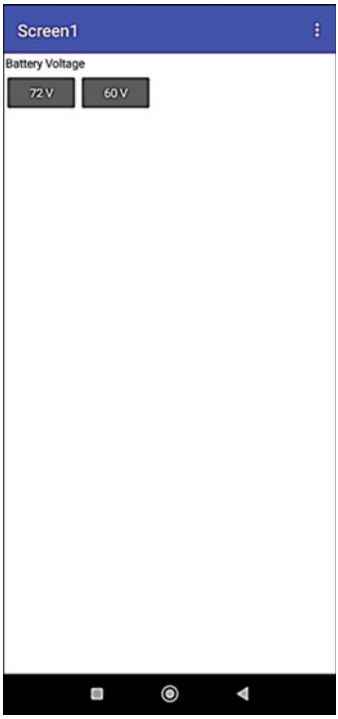


Fig. 5 To remove the battery, the same app will be used to scan the QR code present in the HMI. After scanning, the user will select the desired voltage. Once the battery voltage is selected, the user will be guided by a slot that can be recharged and taken, like the Gesits battery pickup process

5 Conclusion

A thorough analysis of the Battery Swap Station (BSS) in Indonesia and suggested the development of an IoT-based BSS system by applying heterogeneity. A number of factors, including the type of battery and technology used in BSS can be considered in terms of how BSS has developed throughout Indonesia. Based on the above discussions, each battery brand in Indonesia has its own BSS so it can not be used for replacing different voltage batteries. On the prototype we are working on, we have managed to make prototypes of BSS based IoT in Indonesia. The prototype we made can take data from the battery of the two-wheeled motor vehicle of the brand Gesits and the batteries that we have made with close battery voltage (60 and 72 V) of some of the brands on the market in Indonesia. The prototype succeeded by showing in one BSS able to show two different types of battery voltage (heterogeneity) that have several parameters such as battery percentage, estimated battery charging time to full, qr code for payment, availability status, battery voltage, battery temperature, cycle, and battery identity number using Antares as a database where data is stored that is subsequently forwarded to the application and website display in Human Machine Interface (HMI).

References

1. Ministry of Transport Republic of Indonesia. https://dephub.go.id/post/read/tekan-polusi,-kemenhub-dorong-elektrifikasi-kendaraan-bermotor. Accessed 02 Feb 2023
2. Yonghua S, Yang Y, Hu Z (2011) Present status and development trend of batteries for electric vehicles. Power System Technol, 1–7
3. Watts RG (2022) Global warming and the future of the earth, 1st edn. Springer Nature, Colorado
4. OurWorldInData.org. https://ourworldindata.org/co2-and-greenhouse-gas-emissions. Accessed 03 Feb 2023
5. Central Bureau of Statistics Republic of Indonesia. https://www.bps.go.id/indicator/17/57/1/jumlah-kendaraan-bermotor.html. Accessed 02 Feb 2023
6. Main Directorate of Legal Development and Development of State Financial Examination Republic of Indonesia. https://peraturan.bpk.go.id/Home/Details/116973/perpres-no-55-tahun-2019. Accessed 02 Feb 2023
7. Main Directorate of Legal Development and Development of State Financial Examination Republic of Indonesia. ESDM No. 13 Tahun, Berbasis Baterai %5BJDIH BPK RI%5D. https://peraturan.bpk.go.id/Home/Details/163121/permen-esdm-no-13-tahun-2020#:~:text=Permen. Accessed 03 Feb 2023
8. Ministry of Energy and Mineral Resources Republic of Indonesia. https://www.esdm.go.id/en/berita-unit/directorate-general-of-electricity/semakin-bertambah-267-spklu-dan-266-spbklu-telah-dibangun. Accessed 10 Feb 2023
9. Cui D et al (2023) Operation optimization approaches of electric vehicle battery swapping and charging station: a literature review. Energy (263)
10. Zhan W, Wang Z, Zhang L, Liu P, Cui D (2022) A review of siting, sizing, optimal scheduling, and cost-benefit analysis for battery swapping stations. Energy (258)
11. Koirala K, Tamang M, dan Shabbiruddin (2022) Planning and establishment of battery swapping station—a support for faster electric vehicle adoption. J Energy Storage 51
12. Adu-Gyamfi G, Song H, Nketiah E, Obuobi B, Adjei M, dan Cudjoe D (2022) Determinants of adoption intention of battery swap technology for electric vehicles. Energy 251
13. Ministry of Energy and Mineral Resources Republic of Indonesia. https://www.esdm.go.id/id/berita-unit/direktorat-jenderal-ketenagalistrikan/dorong-percepatan-kblbb-pemerintah-tingkatkan-jumlah-spklu-dan-spbklu. Accessed 08 Feb 2023
14. Engineeringforchange. https://www.engineeringforchange.org/solutions/product/gogoro-network-battery-swapping-platform. Accessed 08 Feb 2023
15. PT Kompas Cyber Media. https://otomotif.kompas.com/read/2022/09/14/111200615/spesifikasi-lengkap-motor-listrik-viar-new-q1. Accessed 08 Feb 2023
16. PT Smoot Motor Indonesia Homepage. https://www.smoot.id/tempur. Accessed 08 Feb 2023
17. PT Wika Industri Manufaktur 2022. https://gesitsmotors.com/spesifikasi/. Accessed 08 Feb 2023
18. Volta Indonesia. https://voltaindonesia.com/. Accessed 08 Feb 2023
19. Xiaokang L, Qionghua Z, Kui H, Yuehong S (2007) Battery management system for electric vehicles. Huazhong Univ Sci Tech 35:83–86
20. Ghalkhani M, Habibi S (2023) Review of the Li-ion battery, thermal management, and AI-based battery management system for EV application. Energies 16:1–16
21. Tran MK, Panchal S, Khang TD, Panchal K, Fraser R, Fowler M (2022) Applications of battery management system (BMS) in sustainable transportation: a comprehensive approach from battery modeling to battery integration to the power grid. World Electr Veh J (13):1–11
22. Rodemann T, Kataoka H, Jatschka T, Raidl G, Limmer S, Meguro H (2023) Optimizing the positions of battery swapping stations. Int Electr Veh Technol Conf (6)
23. Sarker MA, Mushfiqur R, Pandzic H, Vazquez O (2013) Electric vehicle battery swapping station: business case and optimization model. Int Conf Connect Veh Expo, 289–294
24. Abd Wahab MH, Anuar NIM, Ambar R, Baharum A, Shanta S, Sulaiman MS, Hanafi HF (2018) IoT-based battery monitoring system for electric vehicle. Int J Eng Technol, 505–510

25. Chaturvedi NA, Klein R, Christensen J, Ahmed J, Kojic A (2010) Algorithms for advanced battery-management systems. Control Syst Mag (30):49–68
26. See KW (2022) Critical review and functional safety of a battery management system for large-scale lithium-ion battery pack technologies. Int J Coal Sci Technol (9)
27. Gabbar HA, Othman AM, Abdussami MR (2021) Review of battery management systems (BMS) development and industrial standards. Technologies (9)
28. Friansa BYK, Haq IN, Santi BM, Kurniadi D, Leksono E (2017) Development of battery monitoring system in smart microgrid based on internet of things (IoT). Eng Phys Int Conf, 482–487
29. Asvika D, Deepika M, Selvabarkavi M, Shalini VB (2023) BMS will be transmitted using internet of things (IoT). Int Conf Intell Comput Control Syst (7)
30. Lohar GV, Kumar MS (2023) Design and implementation of iot based wireless battery management system for electric vehicles. Int J Intell Syst Appl Eng (11)
31. Hommalai C, Khomfoi S (2011) Battery monitoring system by detecting dead battery cells. Int J Sci Res 1:5–15
32. Kale MS, Chaudhari BN (2022) IoT based battery monitoring system. Int Conf Adv Comput Commun Mater
33. Fan R (2012) SMBus quick start guide. Free. Semicond Inc vol Rev (1):1–16
34. CSS Electronics. https://www.csselectronics.com/pages/can-bus-simple-intro-tutorial#what-is-can-bus. Accessed 09 Feb 2023

Business Intelligence in Healthcare: A Review of Knowledge Structures and Level of Analysis

Agung Purnomo, Mega Firdaus, Fairuz Iqbal Maulana, Bigraf Triangga, Muchamad Indung Hikmawan, and Zahra Tazkia Nurul Hikmah

Abstract Business intelligence in healthcare continues to evolve and is essential for increasing healthcare system efficiency, improving patient data management, developing entrepreneurship, and providing deeper insights for healthcare professionals to improve clinical services and outcomes. This study attempts to visually map the sector and examine current business intelligence trends in the healthcare sector. The study analyzes 547 secondary data publications from the Scopus database that were published between 2002 and 2022 using a bibliometric knowledge structure and level of analysis. According to the findings of the research conducted at Universidade do Minho, Portugal, and the United States were the most productive research institutions and countries related to business intelligence in healthcare. This research suggests classifying business intelligence in healthcare studies themes, such as business analytics, information systems, health service, internet of things, and analytics, into BIHIA research themes.

Keywords Bibliometric · Business intelligence · Entrepreneurship · Health business · Healthcare

A. Purnomo (✉) · B. Triangga
Entrepreneurship Department, BINUS Business School Undergraduate Program, Bina Nusantara University, Jakarta 11480, Indonesia
e-mail: agung.purnomo@binus.ac.id

M. Firdaus
Institute for Research and Community Services, Universitas Nahdlatul Ulama Sidoarjo, Sidoarjo 61234, Indonesia

F. I. Maulana
Computer Science Department, Bina Nusantara University, Jakarta 11480, Indonesia

M. I. Hikmawan
Pharmacy Department, Duren Sawit Regional Special Hospital, Jakarta 13430, Indonesia

Z. T. N. Hikmah
Medical Program, Faculty of Medicine, Universitas Islam Malang, Malang 65144, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_35

1 Introduction

Business intelligence (BI) studies have evolved significantly over the years. Initially, BI was limited to simple reporting and data analysis. However, BI has become more complex and comprehensive with advanced technologies and the need for more sophisticated data analysis [1]. Business intelligence (BI) systems integrate data collecting, administration, and storage with analysis instruments to deliver complex knowledge to strategists and decision-makers [2]. BI helps decision-makers by delivering pertinent information presented at the ideal moment, place, and format. BI technologies transform data into usable, accessible information and knowledge through user analysis. BI has become an integral part of modern business strategy [3]. The development of BI has been driven by the need for organizations to gain a competitive edge in the marketplace by leveraging data to identify opportunities and mitigate risks [4]. Business intelligence continues to be developed in the healthcare industry [5, 6].

Business intelligence (BI) systems have been proven to make better roles in the healthcare industry [7]. To manage organizations and operations more effectively, BI systems have become widely used in the healthcare industry [8, 9]. Business intelligence (BI) systems to help hospitals make better decisions in the healthcare industry [10]. The tremendous influence of BI models has been mainly focused on boosting healthcare firms' productivity levels [11]. Now that they have fully developed, BI systems can also be used for epidemiology and health monitoring [11, 12]. Customer segmentation, tailored marketing, and operational assessment are just a few of the marketing data BI systems may offer for the healthcare industry [13]. The distinctive feature of the healthcare sector, which solely places a premium on patient satisfaction, significantly impacts the decision-making of business intelligence [14].

A study on business intelligence in healthcare has been conducted along with business, entrepreneurship, intelligence, and health growth over the past few years. However, Overall, prior studies allied to business intelligence in healthcare are still bound to review a hospital [10], SMEs [15], and countries [16]. There was healthcare research using bibliometrics [17, 18]. There were also studies related to business intelligence using bibliometrics [19, 20]. However, no bibliometric studies have reviewed business intelligence in healthcare, so this is a research gap. There hasn't been a lot of study that looks at healthcare business intelligence from a global perspective, taking into account factors including location, research institutions, time frame, and industry. To add insult to injury, health and business intelligence research rarely thoroughly investigate the association between researchers and keywords. One approach accustomed to looking at the overall picture of research is bibliometrics.

The bibliometric approach can be visualized from the knowledge structure and level of analysis. Numerous researches frequently use bibliometric approaches to present the connection between a field of study and a quantitative approach. Different mathematical and statistical approaches are employed in bibliometrics to measure and assess academic sources based on the knowledge structure and depth of analysis involved [21]. Articles, reviews, reports, conference proceedings, books, and other

materials subjected to peer review are all part of the scope of bibliographic publishing data exploration. Bibliographic analysis from an intellectual structure point of view is a statistical discipline called bibliometrics [22]. This work proposes a research question of visually mapping the landscape and research trends in business intelligence in healthcare. So, this work aims to visually see the current landscape and identify emerging trends in healthcare business intelligence.

It is urgent to conduct business intelligence studies in the healthcare industry using bibliometric methods. The need for healthcare professionals increases with population growth to meet the healthcare necessity of sodality [23]. There is a need to explore technological inventions and their implications in the twenty-first century [24]. Healthcare is one of the most complex and knowledge-demanding fields in the world. Using Business Intelligence (BI) support in healthcare can assist hospitals in better operational management [10].

The following are some of the academic contributions and implications of this research. This study is a novelty since it employs a bibliometric strategy for analyzing business intelligence in healthcare research, integrating analysis of knowledge structure and level of analysis. This method allows for a more all-encompassing comprehension of how business intelligence has been used in the ever-evolving healthcare research field. Furthermore, this study identifies partnership networks between researchers in business intelligence in healthcare publications across a 20-year time period, which contributes to mapping research collaboration. This research is essential to classifying business intelligence in healthcare study themes. This method can lay the groundwork for future studies and assist researchers and practitioners in better comprehending the trajectory and concentration of current studies. This study reveals a community of business intelligence journal authors collaborating on research. Business intelligence in healthcare has implications for the healthcare industry, healthcare practitioners, healthcare entrepreneurs, hospitals, and government health policies.

2 Research Methods

This research employs bibliometric techniques from the standpoint of knowledge structure and level of analysis. The Scopus database was mined for information from 2002 through 2022. The Scopus database's inclusion in this investigation was decided after considering the number of researchers' publications and the research output of their respective universities. This feature also keeps track of the yearly totals for publications, citations, fields of study, and document citations [25]. This study has discovered crucial keywords linked to business intelligence in healthcare research that can be used to search for and locate related publications in the Scopus database globally. The variables and indicators considered in this investigation are displayed in Table 1.

This research uses the terms "healthcare" and "business intelligence" in the Scopus database's title, abstract, and author keywords to extract the desired information. To collect complete public data for all twelve months of each year, data mining was

Table 1 Variables and indicators of research

No	Variables	Indicators
1	Trend research	Number of documents per year
		Number of individual researcher publications
		Number of publications affiliated with research institutions
		Number of document citations
		Number of country publications
2	Mapping research	Network of keywords or publication themes
		Researcher collaboration network

limited to annual data. Five hundred and forty-nine articles were found using the data mining search query option (TITLE-ABS-KEY ("health care" OR healthcare) AND TITLE-ABS-KEY ("business intelligence")) PUBYEAR 2023. The CSV file format was used to retrieve this investigation's Scopus results' metadata.

VOSviewer version 1.6.19 was used to conduct a co-occurrence and co-authorship analysis on the collected documents. This study employs a co-authorship analysis with the author's unit of analysis and a systematic, complete calculating approach with VOSViewer to establish a network of collaborative research among academics [26]. This research generates a keyword map network using VOSViewer and a comprehensive co-occurrence analysis, keyword relation analysis, and a systematic calculation technique. Microsoft Excel created and tabulated simple statistics and tables [27].

3 Research Results and Discussion

An evaluation of business intelligence in healthcare research was carried out by outlining the data outcomes that evolved based on organization and nation affiliations, annual documents, publishing themes, and researchers in business intelligence in healthcare.

3.1 Most Prolific Individual Researchers

About 415 researchers contributed to the work known as business intelligence in healthcare. Ten scholars have published more than the average in business intelligence in healthcare out of a total of 415. Figure 1 shows Abelha, A. has 18 business intelligence in healthcare publications. While Machado, J. has 17 publication documents, Wickramasinghe, N. with 11 documents, Portela, F. with eight documents, Esteves, M. with seven documents, George, J., Guimarães, T., McGlothlin, J.P., Moghimi, F.H., and Peyton, L., have five documents.

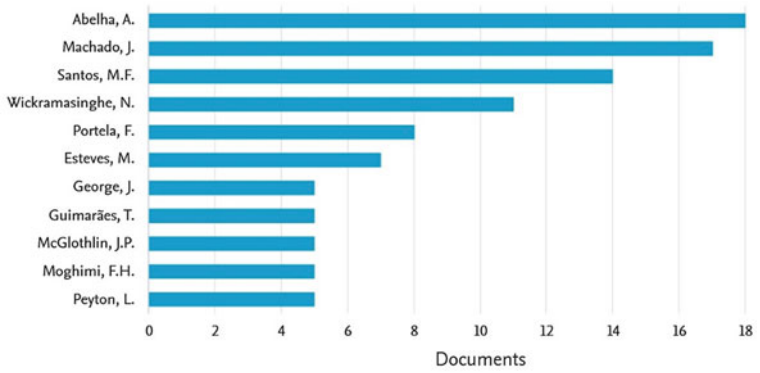


Fig. 1 Most profilic individual researchers

3.2 Most Successful Research Organizations

Business intelligence in healthcare has been the subject of publications from 363 connected academic institutions. As can be seen in Fig. 2, the Universidade do Minho (Portugal) was the most prolific research institution connected to business intelligence in healthcare publishing. This is followed by the University of Ottawa (Canada) with a total of 12 papers, RMIT University (Australia) with a total of 6 documents, Fusion Consulting Inc. (United States), and the University of Colorado Denver (United States), Tampere University (Finland), Cleveland Clinic Foundation (United States) with each institution having five documents. The thirteen groups averaged four publications per member, namely the University of Victoria (Canada), Universidade do Porto (Portugal), Western University (Canada), Instituto Politécnico do Porto (Portugal), University of Toronto (Canada), Macquarie University (Australia), Monash University (Australia), Epworth HealthCare (Australia), Alma Mater Studiorum Università di Bologna (Italy), Iscte–Instituto Universitário de Lisboa (Portugal), Amity University (India), Noorul Islam University (India) and The Faculty of Medicine.

Based on the findings, the Universidade do Minho emerged as the leading center for business intelligence in healthcare research. Universidade do Minho is a university located in Portugal. The university is known for its strong focus on cross-disciplinary research, including information technology and healthcare. In the context of research on business intelligence in healthcare [28].

3.3 Annual Volume of Publication Records

Annual publications showcase the growing use of business intelligence in healthcare research. Business intelligence in healthcare research has been appearing in many

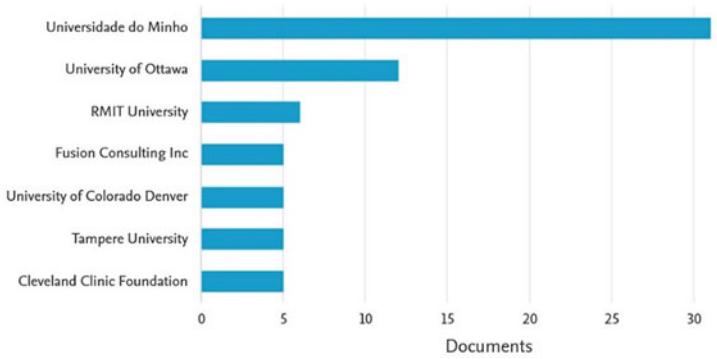


Fig. 2 Most successful research organizations

publications worldwide. Scopus database records show that healthcare business intelligence research started in 2002. According to the annual publishing productivity shown in Fig. 3, business intelligence in healthcare research may be split into four stages: the initial stage, the development stage, the maturity stage, and the growth stage.

- Initial stage (2002–2010)

The first phase ran from 2002 [29] through 2010. At this point, people first started publishing works about business intelligence in the healthcare sector. There were 79 works released between 2002 and 2010 during this time frame. Each publication contained little more than twenty-five papers, and the annual publication rate was around one manuscript.

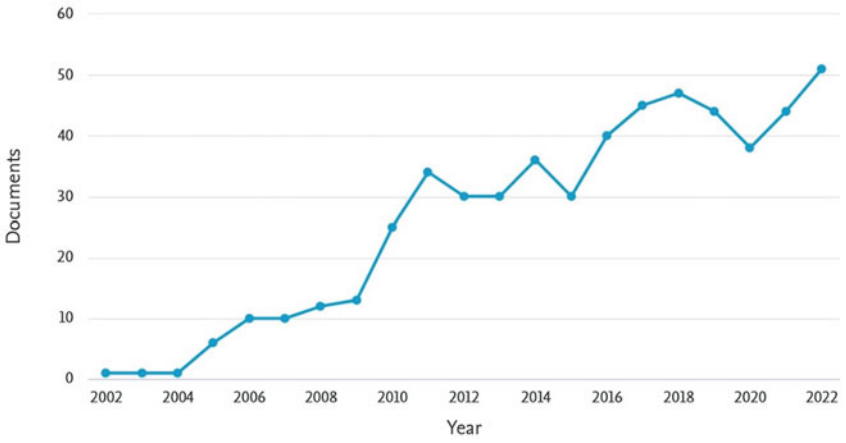


Fig. 3 Annual volume of publication records

- Developmental stage (2011–2014)

The number of drafts increased dramatically as the annual publication rate of 30 papers increased. The annual publishing cap was set at 36. From 2011 to 2014 [30], there was a sharp increase in publication, with 130 papers appearing then. Business intelligence in healthcare research was in its formative years at this time and was gaining prominence.

- Maturity stage (2015–2018)

There is a rise in the quantity of documents from different nations at this maturity level. The number of articles published increased by 162 between 2015 and 2018 [31]. No more than 15 documents were published annually. Ten documents were published annually on average at this point.

- Growth stage (2019–2022)

The average number of documents published annually on healthcare business intelligence is 10. There will be 44 publications in 2019, but just 38 in 2020. In 2021, articles on business intelligence in healthcare increased to 44 papers. With 51 articles, 2022 saw the most significant peak for business intelligence publications. This demonstrates that business intelligence research is growing and has entered a period of rapid expansion.

3.4 Research Theme Map

A research theme map is an overview that uses keyword relationships between publications to find actionable business intelligence in healthcare research. To examine and display the business intelligence included inside the healthcare keyword framework for publication topic maps, the VOSViewer application was employed. The minimum number of papers containing the keywords was set at five. Each cluster needed a minimum of five keywords to be created. This means that out of 3,339 terms, only 259 met the cutoff.

Figure 4 shows five distinct groups of related articles based on the keywords used in business intelligence research. Business intelligence studies can be broken down into five major clusters: Business analytics, Information systems, Health service, Internet of things, and Analytics, abbreviated as BIHIA Research Themes.

a. Business analytics (red)

This business analytics cluster includes keywords such as business analytics, business intelligence (BI), business intelligence platform, business intelligence systems, business process management, and clinical business intelligence. Business Analytics utilizes data collected and processed through BI concepts [32]. The data generated from BI platforms and systems becomes the raw material analyzed by the analytical tools within business analytics [33]. Business analytics leverages the visualization,

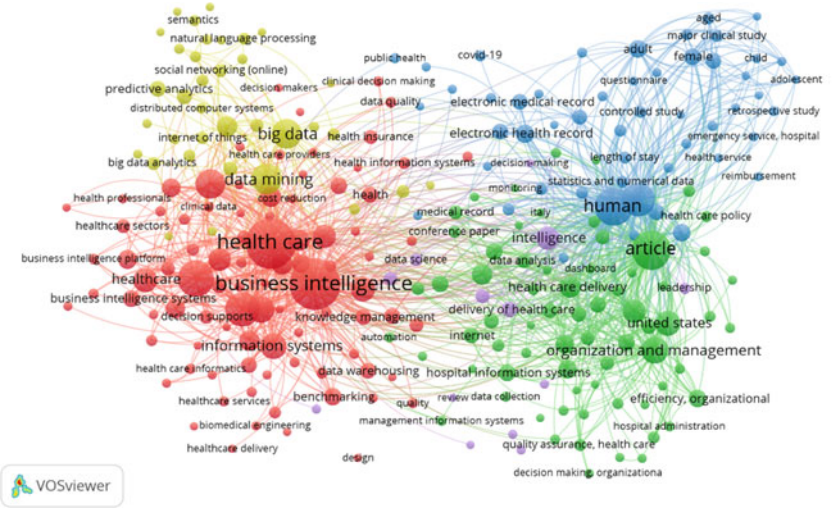


Fig. 4 Research themes map

reporting, and analysis tools built into BI platforms and systems [34]. The BI platform collects, stores, and processes data into reports and dashboards, providing a holistic view of business performance [35].

b. Information system (green)

This information system cluster consists of keywords such as information technology [36], information processing [37], hospital administration, hospital information system [38], and health care system information systems [39] are essential in managing data and information to support decision-making and daily operations.

c. Health service (blue)

Most search terms related to this group of health services involved phrases like health care cost [40], health informatics [41], hospital discharge, hospital emergency service [42], electronic health record [43], emergency service, and hospital [44]. Business intelligence in the health sector can help analyze and understand healthcare costs more deeply.

d. Internet of Things (yellow)

In this group of Internet of Things keywords, researchers uncovered advertising concepts, including big data [45], cloud computing, data mining [46], digital storage, machine learning [47], social media, and social networking (online) [48]. Business intelligence enables healthcare organizations to process, analyze, and gain insights from big data.

e. Analytics (purple)

This analytics cluster includes keywords such as data interpretation [49], statistical, data science [50], decision-making [51], intelligence, and statistical analysis [52]. Business intelligence in healthcare involves the understanding of complex health data. Accurate and in-depth data interpretation helps healthcare professionals make better decisions and plan more effective care.

4 Future Direction and Recommendations

Using solely datasets from the Scopus database has certain limitations for this research. Scopus may have missed some papers dealing with healthcare business intelligence. Several directions for further study in healthcare business intelligence research are suggested by existing literature. The bibliometric analysis of business intelligence in healthcare uses Scopus and Web of Science data. The methodology of the bibliometric study of business intelligence in healthcare is based on the analysis of knowledge structures like social structure and domain levels like sources.

5 Conclusion

The findings reveal a rising pattern of maps and visual patterns in international publications on business intelligence in healthcare from various countries throughout time. With 31 publications in business intelligence in healthcare, the Universidade do Minho (Portugal) stands out as the most prolific research institution worldwide. In 2022, 51 papers were published in the field of business intelligence in healthcare, representing a global high point in this area of research. Business intelligence in healthcare publishing supports a community of researchers collaborating on studies. This study proposes categorizing BIHIA research themes, such as business analytics, information systems, health service, internet of things, and analytics, based on identifying the body of knowledge created for 20 years of scientific publications.

Acknowledgement The authors thank Bina Nusantara University for funding and research ecosystem support.

References

1. Kimball R, Ross M (2013) The data warehouse toolkit: the definitive guide to dimensional modeling. Wiley, New Jersey
2. Reimers M (2014) Leveraging business intelligence to make better decisions: Part I. J Med Pr Manag 29
3. Pozzato N, D'Este L, Gagliazzo L, Vascellari M, Cocchi M, Agnoletti F, Bano L, Barberio A, Dellamaria D, Gobbo F, Schiavon E, Tavella A, Trevisiol K, Viel L, Vio D, Catania S, Vicenzoni

G (2021) Business intelligence tools to optimize the appropriateness of the diagnostic process for clinical and epidemiologic purposes in a multicenter veterinary pathology service. J Vet Diagnostic Investig 33:439–447. https://doi.org/10.1177/10406387211003163
4. Watson HJ, Wixom BH (2007) The current state of business intelligence. Computer (Long Beach Calif) 40:96–99. https://doi.org/10.1109/MC.2007.331
5. Basile LJ, Carbonara N, Pellegrino R, Panniello U (2023) Business intelligence in the healthcare industry: the utilization of a data-driven approach to support clinical decision making. Technovation 120:102482. https://doi.org/10.1016/j.technovation.2022.102482
6. Mboi N, Syailendrawati R, Ostroff SM, Elyazar IR, Glenn SD, Rachmawati T, Nugraheni WP, Ali PB, Trisnantoro L, Adnani QES, Agustiya RI, Laksono AD, Aji B, Amalia L, Ansariadi A, Antriyandarti E, Ardani I, Ariningrum R, Aryastami NK, Djunaedi D, Efendi F, Fauk NK, Ghozali G, Handayani N, Harapan H, Hargono A, Harso AD, Ikawati HD, Indriasih E, Karlina K, Kosen S, Kusnali A, Kusuma D, Lumbantoruan M, Lusiana M, Martini S, Meilinda M, Mubasyiroh R, Musadad DA, Ningrum DNA, Nuraini S, Nurlaela N, Pangaribuan HU, Pawitaningtyas I, Purnomo A, Puspasari HW, Puspasari N, Rahmawaty S, Rangkuti H, Roosihermiatie B, Rosita T, Rustika R, Samodra YL, Samosir JV, Siswanto S, Smith KH, Sudaryanto A, Sugianto S, Suparmi S, Susilarini NK, Tarigan IU, Ticoalu JHV, Trihandini I, Wahidin M, Warouw TS, Widyastuti R, Wulandari RD, Hay SI, Murray CJL, Mokdad AH (2022) The state of health in Indonesia's provinces, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Glob Heal 10:e1632–e1645. https://doi.org/10.1016/S2214-109X(22)00371-0
7. Ur Rehman M, Ullah R, Allowatia H, Perween S, Ul Ain Q, Ammad M, Noorul Hasan T (2023) Elaborating the role of business intelligence (BI) in healthcare management. J Intell Stud Bus 12:26–35. https://doi.org/10.37380/jisib.v12i2.952
8. Jha S, Topol EJ (2016) Adapting to artificial intelligence. JAMA 316:2353. https://doi.org/10.1001/jama.2016.17438
9. Giuffrida MA (2017) Practical application of evidence-based practice. Vet Clin North Am Exot Anim Pract 20:737–748. https://doi.org/10.1016/j.cvex.2017.04.001
10. Kitsios F, Kapetaneas N (2022) Digital transformation in healthcare 4.0: critical factors for business intelligence systems. Information 13:247. https://doi.org/10.3390/info13050247
11. Trevisan G, Linhares LCM, Crim B, Dubey P, Schwartz KJ, Burrough ER, Main RG, Sundberg P, Thurn M, Lages PTF, Corzo CA, Torrison J, Henningson J, Herrman E, Hanzlicek GA, Raghavan R, Marthaler D, Greseth J, Clement T, Christopher-Hennings J, Linhares DCL (2019) Macroepidemiological aspects of porcine reproductive and respiratory syndrome virus detection by major United States veterinary diagnostic laboratories over time, age group, and specimen. PLoS ONE 14:e0223544. https://doi.org/10.1371/journal.pone.0223544
12. Hasler B, Delabouglise A, Babo Martins S (2017) Achieving an optimal allocation of resources for animal health surveillance, intervention and disease mitigation. Rev Sci Tech l'OIE 36:57–66. https://doi.org/10.20506/rst.36.1.2609
13. Kurolov M (2023) Exploring the role of business intelligence systems in digital healthcare marketing: a qualitative study. Int J Soc Sci Res Rev 6:377–383. https://doi.org/10.47814/ijssrr.v6i6.1226
14. Janyapoon S, Liangrokapart J, Tan A (2021) Critical success factors of business intelligence implementation in Thai hospitals. Int J Healthc Inf Syst Informatics 16:1–21. https://doi.org/10.4018/IJHISI.20211001.oa19
15. Mahesh KM, Aithal PS, Sharma KRS (2022) Seven pillars of inclusive ecosystem—transforming healthcare special reference to MSME & SME sectors. Int J Case Stud Business, IT, Educ 6:237–255. https://doi.org/10.47992/IJCSBE.2581.6942.0162
16. Duarte R, Guimarães T, Santos MF (2022) A business intelligence platform for portuguese Misericórdias. Procedia Comput Sci 198:341–346. https://doi.org/10.1016/j.procs.2021.12.251
17. Nti IK, Adekoya AF, Weyori BA, Keyeremeh F (2023) A bibliometric analysis of technology in sustainable healthcare: emerging trends and future directions. Decis Anal J 8:100292. https://doi.org/10.1016/j.dajour.2023.100292

18. Aboelkhir HAB, Elomri A, ElMekkawy TY, Kerbache L, Elakkad MS, Al-Ansari A, Aboumarzouk OM, El Omri A (2022) A bibliometric analysis and visualization of decision support systems for healthcare referral strategies. Int J Environ Res Public Health 19:16952. https://doi.org/10.3390/ijerph192416952
19. de Sousa Lopes B, Amorim V, Au-Yong-Oliveira M, Rua OL (2023) Competitive and business intelligence: a bibliometric analysis. In: Springer proceedings in business and economics. Springer, Cham, pp 187–197. https://doi.org/10.1007/978-3-031-12914-8_15
20. Żółtowski D (2022) Business intelligence in balanced scorecard: bibliometric analysis. Procedia Comput Sci 207:4075–4086. https://doi.org/10.1016/j.procs.2022.09.470
21. Purnomo A, Sari AK, Mufidah E, Asitah N, Aziz A (2020) Digital business: a scientific publication positioning using scientometric analysis. In: 2020 International conference on information management and technology (ICIMTech). IEEE, pp 588–593. https://doi.org/10.1109/ICIMTech50083.2020.9211174
22. IGI Global: What is bibliometric? https://www.igi-global.com/dictionary/bibliometric/49021
23. Grabowska S, Saniuk S (2022) Business models in the industry 4.0 environment—results of web of science bibliometric analysis. J Open Innov Technol Mark Complex 8:19. https://doi.org/10.3390/joitmc8010019
24. Prema RK, Kathiravan M, Shaikh AA (2021) Artificial intelligence in healthcare: 21st century age of rifles—a bibliometric analysis. Asia Pacific J Heal Manag 16:274–281. https://doi.org/10.24083/apjhm.v16i4.1327
25. Purnomo A, Firdaus M, Sutiksno DU, Putra RS, Hasanah U (2021) Mapping of business intelligence research themes: four decade review. In: 2021 IEEE international conference on communication, networks and satellite (COMNETSAT). IEEE, Purwokerto, pp 32–37. https://doi.org/10.1109/COMNETSAT53002.2021.9530790
26. Purnomo A, Rosyidah E, Firdaus M, Asitah N, Septianto A (2020) Data science publication: thirty-six years lesson of scientometric review. In: 2020 International conference on information management and technology (ICIMTech). IEEE, pp 893–898. https://doi.org/10.1109/ICIMTech50083.2020.9211192
27. van Eck NJ, Waltman L (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 84:523–538. https://doi.org/10.1007/s11192-009-0146-3
28. Ferreira J, Portela F, Machado J, Santos MF (2019) Adaptive business intelligence in healthcare—a platform for optimising surgeries. In: Advances in intelligent systems and computing. Springer, Cham, pp 611–620
29. Spil TAM, Stegwee RA, Teitink CJA (2002) Business intelligence in healthcare organizations. In: Proceedings of the 35th annual Hawaii international conference on system sciences. IEEE Computer Society, Hawaii, p 9. https://doi.org/10.1109/HICSS.2002.994108
30. Zheng G, Zhang C, Li L (2014) Bringing business intelligence to health information technology curriculum. J Inf Syst Educ 25:317–325
31. Roth JA, Goebel N, Sakoparnig T, Neubauer S, Kuenzel-Pawlik E, Gerber M, Widmer AF, Abshagen C, Padiyath R, Hug BL, Abshagen C, Fucile G, Gerber M, Goebel N, Hug BL, Jaegle B, Kuenzel-Pawlik E, Neubauer S, Padiyath R, Roth JA, Sakoparnig T, Sengstag T, Spyra D, Widmer AF (2018) Secondary use of routine data in hospitals: description of a scalable analytical platform based on a business intelligence system. JAMIA Open 1:172–177. https://doi.org/10.1093/jamiaopen/ooy039
32. Ettefaghian A, Krause P, De Lusignan S, Stergioulas LK (2022) A business analytics framework for primary care. In: Studies in health technology and informatics. IOS Press, Cardiff. https://doi.org/10.3233/SHTI220926
33. Arefin MS, Hoque MR, Rasul T (2021) Organizational learning culture and business intelligence systems of health-care organizations in an emerging economy. J Knowl Manag 25:573–594. https://doi.org/10.1108/JKM-09-2019-0517
34. Ahmed SYM, Freire SM, Feitosa TMP, Zardo LMG, de Almeida RT (2018) AD-SISCOLO: a decision-support tool to aid the management of a cervical cancer screening program. Res Biomed Eng 34:19–30. https://doi.org/10.1590/2446-4740.08117

35. Namaki Araghi S, Fontanili F, Lamine E, Tancerel L, Benaben F (2018) Applying process mining and RTLS for modeling, and analyzing patients' pathways. In: Proceedings of the 11th international joint conference on biomedical engineering systems and technologies. SCITEPRESS—Science and Technology Publications, Funchal, pp 540–547. https://doi.org/10.5220/0006651605400547
36. Cruz M, Guimarães T, Abelha A, Santos MF (2022) Business analytics components for public health institution—nursing decision area. Procedia Comput Sci 198:347–352. https://doi.org/10.1016/j.procs.2021.12.252
37. Schaeffer C, Booton L, Halleck J, Studeny J, Coustasse A (2017) Big data management in US hospitals. Health Care Manag (Frederick) 36:87–95. https://doi.org/10.1097/HCM.0000000000000139
38. Boland GW, Thrall JH, Duszak R (2015) Business intelligence, data mining, and future trends. J Am Coll Radiol 12:9–11. https://doi.org/10.1016/j.jacr.2014.10.003
39. Basile LJ, Carbonara N, Pellegrino R, Panniello U (2021) The improvement of the clinical decision-making through the business intelligence. In: 2021 29th Mediterranean conference on control and automation (MED). IEEE, Bari, pp 156–161. https://doi.org/10.1109/MED51440.2021.9480240
40. Patton A (2022) Data analytics + business intelligence = operations insights. Oncol Issues 37:36–41. https://doi.org/10.1080/10463356.2022.2083816
41. Ali O, Crvenkovski P, Johnson H (2016) Using a business intelligence data analytics solution in healthcare. In: 2016 IEEE 7th annual information technology, electronics and mobile communication conference (IEMCON). IEEE, Vancouver, pp 1–6. https://doi.org/10.1109/IEMCON.2016.7746328
42. Basis F, Moskovitz K, Tzafrir S (2021) Did the events following the COVID-19 outbreak influence the incidents of violence against hospital staff? Isr J Health Policy Res 10:36. https://doi.org/10.1186/s13584-021-00471-z
43. Kahn RA, Gal JS, Hofer IS, Wax DB, Villar JI, Levin MA (2022) Visual analytics to leverage anesthesia electronic health record. Anesth Analg 135:1057–1063. https://doi.org/10.1213/ANE.0000000000006175
44. Hsu P-C, Huang W-N, Kuo K-M, Yeh Y-T (2022) Application of business intelligence in decision support to hospital management: an example of outpatient clinic schedule arrangement. In: Studies in health technology and informatics. IOS Press, Amsterdam. https://doi.org/10.3233/SHTI220262
45. Mushore R, Kyobe M (2022) Aligning tasks, technology, people, and structures to leverage the value of big data analytics. Procedia Comput Sci 203:739–744. https://doi.org/10.1016/j.procs.2022.07.110
46. Deepa VK, Geetha JRR (2013) Rapid development of applications in data mining. In: 2013 International conference on green high performance computing (ICGHPC). IEEE, pp 1–4. https://doi.org/10.1109/ICGHPC.2013.6533916
47. Patidar PK, Jain R (2022) A review on prediction of diabetes using various machine learning algorithms. ECS Trans 107:5785–5795. https://doi.org/10.1149/10701.5785ecst
48. Farzindar A, Inkpen D (2017) Natural language processing for social media, Second Edition. Synth Lect Hum Lang Technol 10:1–195. https://doi.org/10.2200/S00809ED2V01Y201710HLT038
49. Sousa MJ, Pesqueira AM, Lemos C, Sousa M, Rocha Á (2019) Decision-making based on big data analytics for people management in healthcare organizations. J Med Syst 43:290. https://doi.org/10.1007/s10916-019-1419-x
50. Yee JM, Cross N, Bhargava P (2022) Do-it-yourself business intelligence for the radiologist—lessons learned from 10-year trends in an abdominal imaging division at a tertiary medical center. J Am Coll Radiol 19:329–335. https://doi.org/10.1016/j.jacr.2021.10.007

51. Gupta MK, Hemachandra N, Bhatnagar S (2022) Learning in sequential decision-making under uncertainty. In: Artificial intelligence and machine learning for EDGE computing. Elsevier, New Delhi, pp 75–85. https://doi.org/10.1016/B978-0-12-824054-0.00011-3
52. AbuHalimeh A (2022) Improving data quality in clinical research informatics tools. Front Big Data 5. https://doi.org/10.3389/fdata.2022.871897

Assessing the Effectiveness of Mechanical Sensors for Respiratory Rate Detection

Dyah Titisari and M. Prastawa Assalim T. Putra

Abstract Assessing respiratory rate is pivotal to monitoring an individual's health status. This research endeavors to meticulously evaluate the efficacy of three distinct mechanical sensors: piezoelectric sensors, force-sensitive resistor sensors, and Flex sensors, in the precise measurement of respiratory rates. The experimental design involved the strategic placement of these sensors onto a mannequin's chest cavity, enabling the conversion of chest pressure into tension. Subsequently, a sophisticated data processing system, driven by Arduino technology, was employed, with the results elegantly displayed through a Delphi interface. The outcomes of this extensive experimentation yielded compelling results, with the force-sensitive resistor sensor emerging as the standout performer, boasting a remarkably low error value of 0.2781 ± 0.4481. This exceptional precision was particularly notable when assessing respiratory rates at setting 20. Such findings underscore the pragmatic application of the force-sensitive resistor sensor within healthcare equipment systems, promising user-friendly operation and efficient retrieval of invaluable patient data. This study not only adds depth to our understanding of the suitability of mechanical sensors for respiratory rate monitoring but also accentuates the transformative potential of the force-sensitive resistor sensor in elevating measurement accuracy to unprecedented levels. Furthermore, it lays the foundation for future research endeavors, with promising directions including sensor refinement and the implementation of cutting-edge data processing techniques, aimed at ushering in a new era of excellence in respiratory rate assessment for healthcare applications.

Keywords Respiratory rate · Piezoelectric · Force-sensitive resistor · Flex sensors

D. Titisari · M. P. A. T. Putra (✉)
Department of Medical Elektronic Technonology, Poltekkes Kemenkes Surabaya, Jl. Pucang Jajar Timur No. 10, Surabaya, Indonesia
e-mail: prast77@poltekkesdepkes-sby.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_36

1 Introduction

Respiration, an inherently automatic and life-sustaining process, provides a profound insight into an individual's physiological state [1–4]. The respiratory rate, quantifying the number of breaths taken per minute, emerges as a pivotal parameter in deciphering the complex language of human health. However, the idiosyncrasy of each person's respiratory rate underscores the pressing need for precise and reliable measurement techniques. These techniques not only serve as diagnostic tools for an array of medical conditions but also as the foundation upon which timely healthcare interventions are built [5–8].

Within this landscape, this study embarks on a quest to engineer an effective design for respiratory rate detection [9, 10]. This endeavor hinges on the meticulous evaluation of three distinctive mechanical sensors [11]: piezoelectric sensors, force-sensitive resistor sensors, and Flex sensors. The overarching goal is to discern which sensor type ascends as the paragon of accuracy in quantifying respiratory rates [12]. These mechanical marvels are strategically attached to a belt, which finds its purpose securely nestled in the chest cavity, ingeniously translating nuanced variations in chest pressure into discernible voltage signals [4, 13]. These invaluable data streams are then ushered through a meticulously orchestrated data processing symphony, led by an Arduino system, and are artfully showcased on a Delphi interface, culminating in the presentation of both a respiratory rate signal and a breath count.

While precedent research endeavors have traversed diverse methodologies for respiratory rate measurement, ranging from the employment of ECG sensors to the ingenuity of accelerometers and even venturing into the realm of unconventional techniques such as singing greeting card beepers [14], mechanical sensors present an alluring prospect. They epitomize simplicity, cost-effectiveness, and scalability, all of which are coveted traits in modern healthcare equipment systems. To refine and amplify the pragmatic potential of this approach, the current study meticulously scrutinizes the accuracy of these sensors under controlled conditions, with a judiciously selected mannequin serving as a surrogate for live subjects, ensuring that the experimentation remains precise, repeatable, and controllable.

The advent of three pioneering sensor types—the piezoelectric, the Flex, and the force-sensitive resistor sensors [15, 16]—builds upon the foundation laid by previous research conducted in 2018, which predominantly centered around the application of piezoelectric sensors [17]. This study serves as the natural progression, a systematic exploration that compares and contrasts these multiple sensor types, thereby enriching our collective understanding of the most efficacious mechanical sensor for the faithful quantification of respiratory rates.

In summation, this research endeavor marks a significant stride in the ever-evolving landscape of respiratory rate monitoring [18–20]. It offers a critical assessment, akin to a discerning maestro, meticulously evaluating the accuracy and potential adaptability of mechanical sensors within the intricate symphony of healthcare equipment systems [21]. In the ensuing sections, we descend into the labyrinth of methodology, ascend to the crescendo of results, and engage in the melodious

discourse of discussion, each note unraveling the nuanced performance of these sensors and each phrase echoing with the profound implications they hold for the ever-advancing real of healthcare technology.

2 Method

2.1 Participant Selection

This research was conducted employing a CAE Apollo mannequin as a surrogate for replicating human respiratory patterns. The mannequin's respiratory rate (RR) was systematically adjusted within the range of 10–22 breaths per minute (BRPM), simulating various breathing scenarios. A specialized respiratory belt was meticulously affixed to the left chest area of the mannequin to ensure precise sensor placement and consistency. The sensors—namely, the piezoelectric sensor (PS), force-sensitive resistor sensor (FSRS), and Flex sensor (FS)—were affixed in an alternating fashion on this respiratory belt to acquire data. Each sensor's performance was meticulously evaluated under these controlled conditions, with data collection for each sensor conducted continuously for 2 min at each prescribed respiratory rate setting.

2.2 Data Gathering Approach

The data acquisition process was executed using a well-designed measurement system, depicted in Fig. 1. These sensors, securely positioned in an alternating sequence on the mannequin's chest, generated analog signals that were proportionate to the variations in chest cavity pressure attributed to the act of respiration. To enhance the quality of these analog signals, each sensor was equipped with its dedicated amplifier circuitry. Following this amplification, the analog signals underwent low-pass filtering to effectively eliminate any high-frequency noise components. Subsequently, these filtered signals were amalgamated using an adder circuit to ensure a positively oriented voltage output. These analog signals, now enhanced and optimized, were further subjected to digitization via a microcontroller. At regular one-second intervals, the digital data was transmitted to a connected computer system for subsequent analysis and display.

In this study, data gathering was conducted by measuring the respiratory rate of a mannequin with the aid of three distinct sensors: piezoelectric, force-sensitive resistor, and Flex sensors. The data collection approach involves the following formula:

$$RR = \frac{N}{T} \tag{1}$$

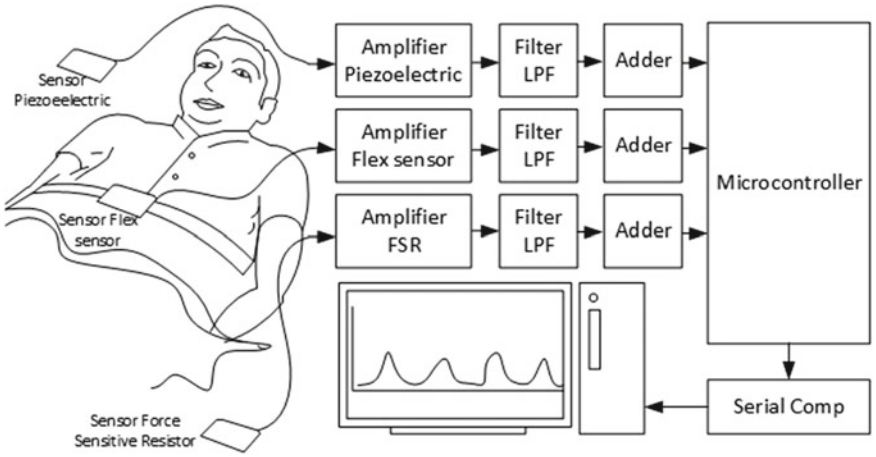


Fig. 1 The picture above shows a block diagram of the measurement system carried out. seen that the sensors used will be alternately installed. The sensor leads will be amplified by an amplifier circuit. Each sensor has its amplifier. After that, it will go through the LPF filter and will enter the adder circuit so that the voltage value is positive. The data is converted into digital data in the microcontroller. Then every 1 s, the data will be sent to the PC

where *RR* represents the respiratory rate in breaths per minute (BPM), N signifies the total number of breaths counted during the data collection period, and *T* denotes the duration of data collection in minutes.

In this context, *RR* is the primary parameter of interest, reflecting the respiratory rate being measured. The total number of breaths (N) is collected by each sensor during a specific data collection period (T), typically lasting for 2 min per setting.

This formula elucidates the fundamental approach undertaken to quantify the respiratory rate using the sensors, providing a clear mathematical foundation for data gathering and subsequent analysis.

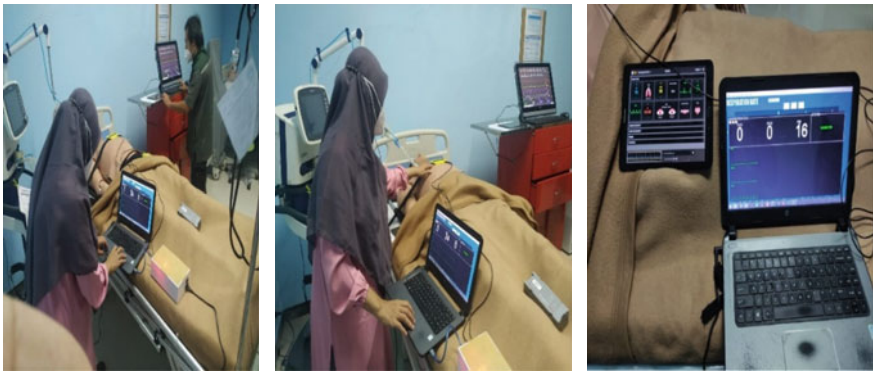
2.3 Data Collection Process

The data collection process was meticulously orchestrated, following a carefully devised sequence of actions. The CAE mannequin was strategically positioned in a supine posture, replicating a typical clinical scenario for patients. The respiratory rate belt, equipped with the three distinct sensors, was expertly positioned beneath the designated area on the mannequin's left chest, precisely as depicted in Fig. 2. To emulate a range of clinical scenarios, the respiratory rate settings were systematically adjusted. This settings from 10 to 22 breaths per minute, with data collection intervals of 2 min for each setting. The entire process was rigorously controlled through a dedicated program running on the mannequin's tablet interface, ensuring utmost accuracy and consistency in respiratory rate settings. The data collection process

is shown in Fig. 3a researcher adjusting the mannequin's respiratory rate value. Figure 3b shows other researchers observing the measurement results produced by one of the sensors. Figure 3c shows a direct comparison of the results produced by one of the sensors on a PC with the setting value on the mannequin's tablet.



Fig. 2 You can see the procedure for installing sensors in turn on the CAE Apollo mannequin. The sensor will be placed under the yellow box located on the left chest of the mannequin



(a) **(b)** **(c)**

Fig. 3 The data collection process is shown in **a** a researcher adjusting the mannequin's respiratory rate value. **b** shows other researchers observing the measurement results produced by one of the sensors. **c** shows a direct comparison of the results produced by one of the sensors on a PC with the setting value on the mannequin's tablet

2.4 Data Processing and Analysis

The acquired dataset encompassed both the respiratory rate signal and the waveform shape for each sensor at various respiratory rate settings. The research effort yielded a substantial volume of data, with a total of 4287 respiratory rate data points gathered for each sensor during a single data collection session. To ascertain the accuracy of the sensors, these data points were meticulously compared against the predetermined respiratory rate settings to calculate error values. The error values were expressed as percentages to establish a standardized measure of accuracy. To deepen the understanding, the error values were categorized into two distinct ranges: the first range encompassed respiratory rate settings ranging from 10 to 16 breaths per minute, while the second range spanned settings from 18 to 22 breaths per minute. This categorization allowed for a finer-grained assessment of sensor accuracy across differing respiratory rate scenarios.

Statistical Evaluation. The statistical analysis of the dataset involved several key steps, including the computation of means (μ) and standard deviations (σ) for the error values associated with each sensor. Moreover, an analysis of variance (ANOVA) was meticulously executed to gauge the significance of any differences in accuracy observed among the sensors at different respiratory rate settings. The formula for ANOVA is as follows:

$$F = \frac{MSB}{MSW} \tag{2}$$

where, F represents the ANOVA test statistic, MSB denotes the Mean Square Between (variance between group means), MSW signifies the Mean Square Within (variance within individual groups or treatments), The calculations involve the following steps:

Determine the P-value. The calculated F-statistic is compared to the F-distribution with dfB and dfW degrees of freedom to determine the p-value, which represents the probability of obtaining an F-statistic as extreme as the one calculated from the data, assuming no significant differences exist between the groups.

The significance level (alpha, typically set at 0.05) is used to make a decision about whether there are statistically significant differences among the sensors' accuracy at different respiratory rate settings. If the p-value is less than alpha, it suggests that at least one sensor's accuracy significantly differs from the others, warranting further investigation.

This rigorous statistical analysis enhances the depth and validity of the study's conclusions regarding the comparative accuracy of the mechanical sensors in respiratory rate detection across various settings.

3 Results

In this section, we delve into the findings of our study, which aimed to evaluate the accuracy and effectiveness of three distinct mechanical sensors: piezoelectric sensors, force-sensitive resistor sensors, and Flex sensors in measuring respiratory rates. The results are presented systematically, encompassing the display of respiratory rate signals, error value calculations, statistical evaluations, comparative analyses, and their practical implications.

3.1 Respiratory Rate Signal Display

The purpose of our study was to develop an effective respiratory rate detection system by assessing these mechanical sensors' performance. Figure 4 showcases the results of our measurements using one of the sensors. It displays the respiratory rate signal's form and values on a PC, with similar results observed when using other sensors.

To comprehensively assess the sensors' accuracy, we conducted measurements across a range of respiratory rate settings (from 10 to 22 breaths per minute, with 2 breaths per minute increments). For each sensor at each setting, we collected data for 2 min. This resulted in a substantial dataset, from which we calculated error values by comparing the measured respiratory rates with the preset values.

Error Values at 10–16 Settings per Minute in Fig. 5. In this range of respiratory rate settings (10–16 breaths per minute), we observed varying error values for each sensor. For example, when the respiratory rate was set at 16 breaths per minute, the error values were as follows: Piezoelectric sensor: 0.8193 ± StDev 0.3847, Force-sensitive resistor sensor: 0.6389 ± StDev 0.4803, Flex sensor: 1.4056 ± StDev 3.0433.



Fig. 4 The PC display when measuring the respiratory rate using the force sensitive resistor sensor at the setting. The mannequin's respiratory rate value is 12 per minute. There is shown the value and form of the signal

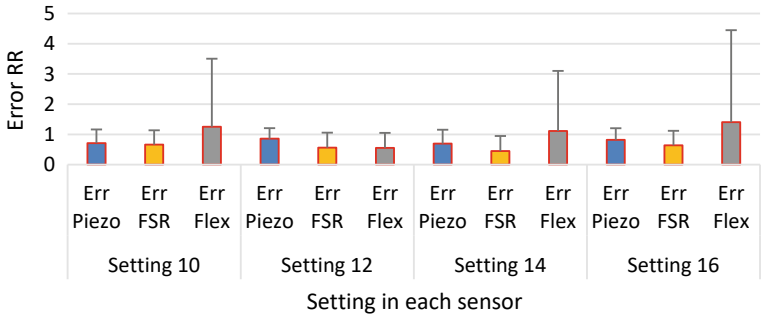


Fig. 5 The PC Illustration of data processing comparison of mannequin setting values and measurement results for each sensor at 10–16 settings per minute. the vertical axis is the percentage of error. shows the error result in setting the RR (Respiration Rate) value of 10–16

Figure 6 shows the error values at the 18–22 breaths per minute setting. In this range of breathing rate settings (18–22 breaths per minute), there are varying error values for each sensor. For example, when the breathing rate is set at 20 breaths per minute, the error values are as follows.

When setting the RR 20 value, the error value on the piezoelectric sensor is 0.3444 ± StDev 0.4752; error value on force sensitive resistor sensors is 0.2781 ± StDev 0.4481 and the error value on Flex sensor is 1.9064 ± StDev 4.0557. So, the force-sensitive resistor sensor exhibited the smallest error value (0.2781 ± 0.4481) when measuring the respiratory rate at setting 20.

To validate the hypothesis proposed against the observed data, whether it is in accordance with the research results or not, the statistical value of P Value is required. P value is a statistical measure that helps determine whether the hypothesis is true or not. P value is a number that lies between 0 and 1.

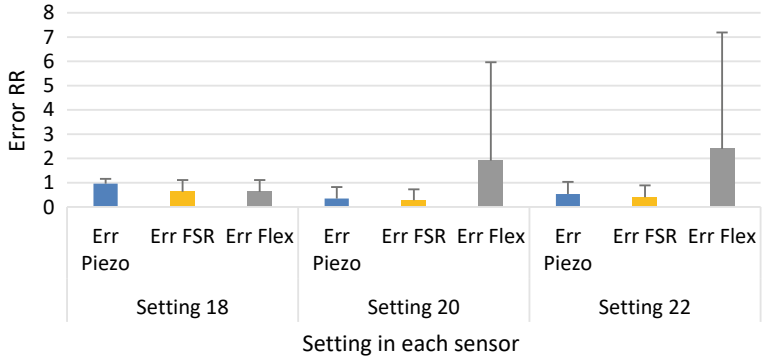


Fig. 6 Illustration of data processing comparison of mannequin setting values and measurement results for each sensor at 18–22 settings per minute. the vertical axis is the percentage of error. shows the error result in setting the RR (Respiration Rate) value of 18–22

Table 1 Summary of the accuracy of each sensor at different RR settings with alpha value = 0.05

Setting	p-value
10	0
12	0
14	0
16	0
18	0
20	0

The significance level (α) is a predetermined threshold that must be set by the researcher, generally set as 0.05. Table shows that in the setting of 10–20 per minute with an alpha setting of 0.05, all pvalue results are zero. So when the p-value is below the selected alpha value, then we say the test results are statistically significant (see Table 1).

4 Discussion

The findings of this study shed light on the effectiveness of mechanical sensors, including piezoelectric sensors, force-sensitive resistor sensors, and Flex sensors, in measuring respiratory rates. The discussion encompasses a comprehensive analysis of the results, a comparison with prior research, and insights into the practical implications of the study.

4.1 Comparative Analysis of Sensor Accuracy

The primary objective of this research was to assess the accuracy of three distinct mechanical sensors in determining respiratory rates. The data revealed variations in sensor performance, particularly in terms of error values at different respiratory rate settings. Notably, the force-sensitive resistor sensor exhibited the smallest error value (0.2781 ± 0.4481) when measuring the respiratory rate at setting 20. This result suggests that the force-sensitive resistor sensor outperformed the other sensors in this specific scenario. However, it is crucial to acknowledge that the accuracy of these sensors was subject to fluctuations across different respiratory rate settings. The flex sensor, in particular, demonstrated a more significant error at various settings, with an error value of $1.9064 \pm$ StDev 4.0557 at setting 20. This discrepancy highlights the sensitivity of the flex sensor, which might be attributed to its design and placement.

4.2 Comparison with Prior Research

To contextualize the study's findings, a comparison with previous research is essential. Prior studies have explored various methods for respiratory rate measurement, including ECG sensors, accelerometers, and piezoelectric sensors. For instance, research by Bhattacharya et al. [19] utilized a flexible sensor and reported an average error of 7.9. These variations in error rates emphasize the significance of sensor selection and calibration in respiratory rate monitoring. Moreover, the study aligns with the work of Karhu [15], which employed a breathing belt transducer using a singing greeting card beeper. While this method generated a breathing chart directly from the sensor, the current research delves deeper into the accuracy of mechanical sensors, providing valuable insights for potential applications in healthcare and diagnostics [22].

4.3 Practical Implications

The results of this study hold practical implications for healthcare equipment systems and patient data retrieval. The identification of the force-sensitive resistor sensor as the most accurate at setting 20 suggests its potential utility in scenarios requiring precise respiratory rate monitoring. However, it is imperative to consider the specific context and requirements of each application when selecting an appropriate sensor.

The study underscores the importance of sensor sensitivity and placement, as these factors significantly impact accuracy. Further research and sensor refinement may lead to improved accuracy across a broader range of settings.

4.4 Limitations and Future Directions

While this study provides valuable insights into sensor accuracy, it is essential to acknowledge its limitations. The use of a mannequin as a breathing simulator may not fully capture the complexity of real-world respiratory rate measurement in clinical settings. Additionally, sensor placement was not rigorously evaluated, potentially introducing bias into the results.

Future research could address these limitations by conducting experiments with human subjects and optimizing sensor placement. Furthermore, sensor design and digital filtering techniques could be refined to enhance accuracy and reliability.

5 Conclusion

In conclusion, this study aimed to develop an effective design for respiration rate detection by evaluating the performance of three different mechanical sensors: piezoelectric sensors, force-sensitive resistor sensors, and flex sensors. The respiratory rate, a vital physiological indicator of a patient's health, plays a crucial role in diagnosing diseases and assessing overall well-being.

In conclusion, this study aimed to develop an effective design for respiration rate detection by evaluating the performance of three different mechanical sensors: piezoelectric sensors, force-sensitive resistor sensors, and flex sensors. The respiratory rate, a vital physiological indicator of a patient's health, plays a crucial role in diagnosing diseases and assessing overall well-being.

To further advance this research, future studies could focus on enhancing sensor design, implementing more precise digital filtering techniques, and conducting experiments with live human subjects to better reflect real-world scenarios.

Ultimately, the findings of this study contribute to the ongoing efforts to develop innovative and accurate respiratory monitoring systems. These advancements hold great potential for application in healthcare equipment systems, enabling simplified and efficient monitoring of patients' respiratory rates, and facilitating data retrieval for improved patient care and diagnosis.

References

1. Bao X, Howard M, Niazi IK, Kamavuako EN (2020) Comparison between embroidered and gel electrodes on ECG-derived comparison between embroidered and gel electrodes on ECG-derived respiration rate, no July. IEEE, pp 1–5. https://doi.org/10.1109/EMBC44109.2020.9176485
2. Eisenberg ME, Givony D, Levin R (2020) Acoustic respiration rate and pulse oximetry-derived respiration rate: a clinical comparison study. J Clin Monit Comput 34(1):139–146. https://doi.org/10.1007/s10877-018-0222-4
3. Panahi A, Hassanzadeh A, Moulavi A (2020) Design of a low cost, double triangle, piezoelectric sensor for respiratory monitoring applications. Sens Bio-Sens Res 30(September):100378. https://doi.org/10.1016/j.sbsr.2020.100378
4. Lepine NN, Tajima T, Ogasawara T, Kasahara R, Koizumi H (2016) Robust respiration rate estimation using adaptive Kalman filtering with textile ECG sensor and accelerometer. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society EMBS, vol 2016-Octob, pp 3797–3800. https://doi.org/10.1109/EMBC.2016.7591555
5. Yearmwar MBB, Mhetre PM (2018) Respiration rate measurement: a different approach. Int J Pure Appl Math 118(24):1–9
6. Gede ID, Utomo B, Amrinsani F, Purwanto E (2021) Low cost health monitoring system based on internet of things using email notification. Indones J Electron Electromed Med Informatics 3(2):45–52
7. Sorie W, Kamara N (2020) Smart sensors and their application in monitoring respiratory rate
8. Andurkar AGANESH (2017) Real time respiration rate measurement using temperature sensor. Int J Recent Innov Trends Comput Commun 5(6):605–607

9. Saatchi R, Burke D, Elphick HE (2011) Respiration rate monitoring methods: a review. Pediatr Pulmonol (July 2020):1–48. https://doi.org/10.1002/ppul.21416
10. Shankar N, Sankar K (2018) Measurement of respiratory rate using peizoelectric sensor. Int J Sci Res Rev 7(1):184–188
11. Dash S, Shelley KH, Silverman DG, Chon KH, Member S (2010) Estimation of respiratory rate from ECG, photoplethysmogram, and piezoelectric pulse transducer signals: a comparative study of time—frequency methods. IEEE Trans Biomed Eng 57(5):1099–1107
12. Ps KB, Jatti A (2015) Respiration and heart rate monitoring from photoplethysmograph signal. J Dep Instrum 3(1):8–16. https://doi.org/10.13189/asp.2015.030102
13. Zhang Q, Chen X, Zhan Q, Yang T, Xia S (2017) Respiration-based emotion recognition with deep learning. Comput Ind 92:84–90
14. Bhaskar A, Subramani S, Ojha R (2014) Respiratory belt transducer constructed using a singing greeting card beeper. J Am Physiol Soc (March 2013):2012–2014. https://doi.org/10.1152/advan.00166.2012
15. Karhu T (2018) Piezoelectric sensors for measuring biosignals. J J Chem Pharm Sci 7:92
16. Klap T, Shinar Z (2018) Using piezoelectric sensor for continuous-contact-free monitoring of heart and respiration rates in real-life hospital settings using piezoelectric sensor for continuous-contact-free monitoring of heart and respiration rates in real-life hospital settings. J Comput Cardiol 40(February):671–674
17. Ardiyanto A, Yulianto E, Titisari D (2018) Sabuk respiration rate dengan sensor piezoelektrik. Indones J Electron Electromed Med Informatics, 1–10
18. Motin MA, Karmakar CK, Palaniswami M (2016) An EEMD-PCA approach to extract heart rate, respiratory rate and respiratory activity from PPG signal. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society (EMBS), vol 2016-Octob, pp 3817–3820. https://doi.org/10.1109/EMBC.2016.7591560
19. Bhattacharya R, Bandyopadhyay N, Kalaivani S (2017) Real time android app based respiration rate monitor. In: International conference on electronics, communication and aerospace technology, pp 709–712
20. Gede ID, Wisana H, Triwiyanto T, Luthfiyah S (2019) Measuring respiration rate based on android. J Elektron Elektromedis dan Inform Medis Indones 1(1):20–26. https://doi.org/10.35882/ijeeemi.v1i1.4
21. Bhattacharya R, Bandyopadhyay N, Kalaivani S (2017) Real time Android app based respiration rate monitor. In: 2017 international conference of electronics, communication and aerospace technology (ICECA), vol 1, pp 709–712
22. Vanegas E (2021) The effect of measurement trends in belt breathing sensors. Eng Proc 4–8

Obesity Prediction Approach Based Habit Parameter and Clinical Variable Using Self Organizing Map

Lilik Anifah, Haryanto, I. G. P Asto Buditjahjanto, R. R. Hapsari Peni Agustin Tjahyaningtijas, and Lusia Rakhmawati

Abstract Obesity is a health problem in the twenty-first century, and can trigger several dangerous diseases. Several studies have been conducted in predicting obesity rates. This study aims to propose an obesity prediction approach using a Self Organizing Map by considering habit parameters and clinical variables as input from the system. The variables considered are gender, age, height, weight, family history with overweight, FAVC, FCVC, NCP, CAEC, whether smoking or not, CH2O, SCC, FAF, TUE, CALC, and transportation information used every day. In this study, the level of obesity was divided into 7 cluster levels, namely insufficient weight, normal weight, overweight level I, overweight level II, obesity type I, obesity type II, and obesity type III. The stages of this research are data preprocessing, initialization process, learning process, and testing process. The learning process is carried out using a Self Organizing Map. System input consists of 16 parameters with 1 obesity level system output. The learning and testing process has been carried out and the results of the first testing process were 81.4%, the precision value was 0.825, while the recall value was 0.803. The second testing process obtained a system accuracy value of 74.3%, the precision value was 0.753, while the recall value was 0.743. However, this system has been able to classify according to the cluster well. The contribution of this research is that it can be a decision support system in predicting obesity levels, and based on the variables considered users can improve their living habits to avoid obesity and live healthier in the future.

Keywords Obesity · Prediction approach · Habit · Clinical variable · SOM

L. Anifah (✉) · I. G. P. Asto Buditjahjanto · R. R. H. P. A. Tjahyaningtijas · L. Rakhmawati
Department of Electrical Engineering, Faculty of Engineering, Universitas Negeri Surabaya, Kampus Unesa Ketintang Surabaya, Surabaya, Indonesia
e-mail: lilikanifah@unesa.ac.id

Haryanto
Department of Electrical Engineering, Faculty of Engineering, Universitas Trunojoyo Madura, Kampus Trunojoyo Madura, Telang Bangkalan Madura, Bangkalan, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_37

1 Introduction

Obesity can be defined as being overweight for a certain height, and currently remains a global health problem [1]. Obesity is a complex condition, one of which is caused by a mismatch between intake and output, which is the main mechanism contributing to weight gain [2]. Obesity can also be defined as a condition in which excessive accumulation of abnormal fat in adipose tissue can potentially interfere with one's health [3]. It is unfortunate that in recent decades the prevalence of overweight and obesity has increased dramatically [4].

Obesity is a health problem in the twenty-first century, and can trigger several diseases such as digestive disorders, type 2 diabetes mellitus, cardiorespiratory dysfunction, and cancer. Several factors that cause obesity include genetic, physiological, behavioral, and environmental factors [5]. Based on other studies, it is stated that the causes of obesity are intrinsic and extrinsic (some call them contributors). Intrinsic causes are conditions where there is a disturbance in the regulation of energy balance due to defects in physiology or behavior that affects excessive energy storage. This is influenced by genes that can affect brain function and behavior, so that genes may have a role in the development of obesity. While extrinsic factors (contributors) are factors such as food consumed, and also physical activity carried out by individuals [6]. Several studies have suggested that genes have an effect on obesity, including 17 genes that encode peptides that send signals of one's hunger and satiety, besides that genes are also involved in adipocyte growth and differentiation, and genes involved in the control of energy expenditure [7].

Obesity can lead to a higher risk of premature death, as well as several chronic diseases such as coronary artery disease, diabetes, stroke, hypertension, osteoarthritis, obstructive respiratory disease and sleep apnea, and gallstones. This is due to the mass effect of expanding adipose tissue or metabolic effects that occur. Obesity is also associated with several types of cancer. The impact on the national economy is the high cost of medical care [8]. Obesity also causes several conditions in the cardiovascular system, neurologic system, coronary heart disease, idiopathic intracranial, hypertension, meralgia paresthetica, pulmonary embolism, stroke, varicose veins, psychosocial, gastrointestinal system, depression, cholelithiasis, social/employment discrimination, gastroesophageal reflux disease (GERD), work disability, colon cancer, hepatic steatosis, reproductive/endocrine systems, hernias, amenorrhea, nonalcoholic steatohepatitis (NASH), breast cancer, cushing syndrome, integumental system, type 2 diabetes, cellulitis, dyslipidemia, carbuncles, glucose intolerance, hygiene problems, hypothyroidism, intertrigo, infertility, venous, stasis of legs, insulin resistance, uterine cancer, musculoskeletal system, polycystic ovary syndrome (PCOS), immobility, lower back pain, respiratory system, osteoarthritis dyspnea and prostate cancer [9].

Research on identifying obesity has been carried out by several researchers, one of which is identifying risk factors that predict obesity using a machine learning classifier and analyzing the best accuracy among the classifiers used to predict obesity. The findings obtained were that there were four features that most influenced all sex

and age groups in predicting obesity, namely triglycerides, glycated hemoglobin, ALT (SGPT), and uric acid. Our findings show that obesity risk factors are sensitive to age and gender under different machine learning algorithms. The highest performance for the age group of both sexes is 70% [10]. A predictive model for identifying obesity levels has also been developed using a decision tree, support vector machine, k-nearest neighbor, gaussian naive bayes, multilayer perceptron, random forest, gradient boosting, and extreme gradient boosting. The input data is the physical condition and how a person's eating habits are. The results obtained are 78% accuracy, 79% precision, 78% recall, and 78% F1 score [11]. Another study predicted obesity for children at age five based on data from one healthy child visit, multiple healthy child visits under the age of two, and a random number of healthy child visits under the age of five. The input data used are BMI, gestational age, birth BMI, birth height, birth weight, and gender. The results show the best accuracy calculation is 89% [12]. A predictive model for obesity in 10 year olds has also been developed. The data used as input is the child's data and also the mother's, the child's data used is gender, eating habits, activity, and previous body mass index. While the mother's data which is used as input from the system is the level of education, self-esteem, and body mass index. The results of the experiment showed that the Area Under the Receiver Operator Characteristic Curve value was 0.82 and an accuracy value of 76% [13].

Several studies to predict obesity rates have been carried out. However, these studies mention the various variables that are considered as system input variables. Research [10] factors that become system inputs are triglycerides, glycated hemoglobin, ALT (SGPT), and uric acid, while physical condition and eating habits [11]. The age factor is also a limitation for these studies, some predict for adults, as well as children.

The obesity prediction methods that have been applied are decision trees, support vector machines, k-nearest neighbors, Gaussian naive Bayes, multilayer perceptron, random forest, gradient boosting, and extreme gradient boosting, but however these methods have been applied, the input data used is different. This research uses a Self Organizing Map, where this method has never been applied before. The difference from the research above is the method used in predicting. This study aims to propose an obesity prediction approach using a Self Organizing Map by considering habit parameters and clinical variables as input from the system. The contribution of this research is that it can be a decision support system in predicting obesity levels, and based on the variables considered users can improve their living habits to avoid obesity and live healthier in the future.

2 Method

2.1 Data and Material

The data used in this study were obtained from the UCI Machine Learning Repository which consists of 17 attributes that are considered in determining the level of obesity [14]. This data was obtained from research on estimating obesity levels based on eating habits and physical conditions from Colombia, Peru and Mexico [15]. The variables considered are gender, age, height, weight, family history with overweight, Frequent consumption of high caloric food (FAVC), frequency of consumption of vegetables (FCVC), Number of main meals (NCP), Consumption of food between meals CAEC, whether smoking or not, Consumption of water daily (CH2O), calories monitoring (SCC), frequency of physical activity (FAF), technological devices used (TUE), frequency of alcohol consumed (CALC), and transportation information used every day. In this study, the level of obesity was divided into 7 cluster levels, namely insufficient weight, normal weight, overweight level I, overweight level II, obesity type I, obesity type II, and obesity type III.

Data is divided into 2 types, namely data for learning and data for testing process. This learning data is represented by each cluster of 10 data, so that the learning data is 70 data. While the testing data is 2041 data consisting of 7 clusters of obesity levels.

2.2 Research Stage

The stages of this research are data preprocessing, initialization process, learning process, and testing process. This preprocessing process is to ensure that the data used is complete without any blank data. This is because when the learning process must be ensured that all data variables have value. The preprocessing process also processes data which is text data into numeric data, where grading text data can be represented with a (numeric) value so that it can be processed during the learning process.

The initialization stage is parameter initialization and weight initialization. Parameter initialization includes determining the maximum number of epochs, learning rate values, and minimum errors. The initialization of the weights is to determine the initial value of the weights of each cluster during the learning process.

The learning process is carried out using a Self Organizing Map. System input consists of 16 parameters (which have been described in Sect. 2.1) with 1 obesity level system output. The output system is divided into 7 clusters of obesity levels, namely insufficient weight, normal weight, overweight level I, overweight level II, obesity type I, obesity type II, and obesity type III. The result of this learning process is to obtain a weight value that will be used in the testing process. The testing process

is by recalling the weights that have been generated in the learning process. Based on the results of the testing process the system performance will be calculated.

2.3 Algorithm

The self-organizing has undergone improvements, one of which is growing hierarchical self-organizing map [16], Convolutional Self-Organizing Map [17], HSOM to visualize Self-Organizing Maps to categorize data [18], and Deep Convolutional Self-Organizing Map Network [19]. This method has also been applied in various fields, for example medicine [20], education [21], smart farming [22], and other applications.

The method used is Self Organizing Map with the following algorithm [23]:

- Initialize the input vector.
- Determined the target of each input.
- Determined the initial weight (gold standard) for each cluster.
- Determine the learning rate value where the value is between 0.1 and 0.9.
- Repeat steps 6–9 until it reaches the specified epoch value.
- The training process starts with data learning.
- Find the shortest distance from each cluster using the eucledian distance formula (formula 1). This process is sought for the lowest distance, hereinafter referred to as winning neurons [24].

$$dx = \sum_{i=1}^{i=13} \left| (xi_i - xgd_i) \right| \tag{1}$$

where:

dx is eucledian distance

xi_i is input data

xgd_i is gold satandard.

- Comparing the results of the shortest distance to each cluster, if the calculated cluster results match the specified target then the weight is fixed. If it is different, a new weight calculation is carried out as in formula 2.

$$w_ij\,(t+1) = w_ij\,(t) + \propto (t)\left[x_ij - w_ij\,(t)\right] \tag{2}$$

- Perform weight updates.
- Save new weights.

The learning process uses a learning rate of 0.1 and a number of epochs of 1000.

2.4 System Evaluation

The results of the testing process are then analyzed for system performance by calculating accuracy, precision, and recall/sensitivity. Calculation of accuracy, precision, and recall/sensitivity as in formulas (3)–(5) [25, 26].

$$acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

where:

Acc is accuration

TP is True Possitive

TN is True Negative

FP is True Possitive

FN is False Negative.

3 Results

The variable analysis for each cluster is described in Tables 1, 2 and 3. Table 1 describes the minimum value for each variable, Table 2 describes the maximum value for each variable, and Table 3 describes the average value for each cluster variable 1–7.

After conducting an analysis of the minimum, maximum, and average values of each variable in each class, then the learning process is carried out. This learning process uses 10 data in each cluster. In the learning process, the selection of the gold standard for each cluster is also carried out. The selection of this gold standard is done randomly, it is expected that the gold standard chosen can represent each cluster. Illustration of the gold standard data variable values for each cluster is shown in Fig. 1.

Based on Fig. 1 then it cannot be modeled linearly, this is because several variables have almost the same values. After selecting the gold standard for each cluster, learning is then carried out with the parameter number of epochs of 1000 and the learning rate used is 0.1.

The testing process is carried out by recalling the weights generated from the learning process. The testing process is carried out in 2 stages, the first stage uses data in the learning process, and the second stage uses data in the testing process. This

Table 1 Minimum value of cluster 1–7 for each variable

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Age (year)	18,00	20,00	19,77	22,00	19,85	26,84	18,31
Height (m)	1,74	1,56	1,54	1,69	1,79	1,69	1,56
Weight (kg)	50,25	45,00	64,73	80,41	104,19	102,60	102,00
Family history with overweight (1: yes; 2: no)	1,00	1,00	2,00	1,00	1,00	1,00	1,00
FAVC (1: yes; 2: no)	1,00	2,00	1,00	1,00	1,00	1,00	1,00
FCVC (1: never; 2: sometimes; 3: always)	1,87	2,00	2,71	2,00	2,00	1,00	3,00
NCP (times)	3,00	3,00	2,67	1,00	3,00	1,09	3,00
CAEC (1: no; 2: sometimes; 3: frequenly; 4: always)	2,00	2,00	2,00	2,00	2,00	1,00	2,00
SMOKE (1: yes; 2: no)	1,00	1,00	1,00	1,00	1,00	1,00	1,00
CH2O (liter)	1,28	2,00	1,98	1,03	2,15	1,37	1,00
SCC (1: yes; 2: no)	1,00	1,00	1,00	1,00	1,00	1,00	1,00
FAF (times a week)	0,68	0,00	0,45	0,00	1,61	0,00	0,00
TUE (hours)	1,49	1,00	0,00	0,00	0,00	0,00	0,71
CALC (1: no; 2: sometimes; 3: frequenly; 4: always)	1,00	1,00	2,00	1,00	1,00	1,00	2,00
MTRANS (1: Automobile; 2: Motorbike; 3: Bike; 4: Public Transportation; 5: Walking)	1,00	1,00	1,00	1,00	1,00	1,00	1,00

is done in order to analyze to what extent the learning process recognizes learning data. Meanwhile, using learning data is intended to analyze the extent to which the designed system can recognize data other than learning data.

The results of testing using learning data are described in Tables 4 and 5. Table 5 is confusion matrix, which is based on this the system performance is capable to be analyzed. Based on confusion matrix, it can also be analyzed to what extent the system can classify data when compared to the actual data. The diagonal part of the confusion matrix explains the correct data which is classified correctly by the system or known as True Positive (TP). Cluster 1 data that are classified correctly have 8 data, while those that are classified incorrectly have 2 data, which should be classified as cluster 1 but classified as cluster 2.

Calculation of the performance of the system by calculating the value of accuracy, precision, and recall using formulas (3)–(5). The accuracy produced in the first testing process (Table 5) is 81.4%, the precision value is 0.825, while the recall value is 0.803.

Table 2 Maximum value of cluster 1–7 for each variable

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Age (year)	23,00	22,00	23,00	23,00	38,00	41,00	23,66
Height (m)	1,89	1,88	1,67	1,75	1,75	1,84	1,75
Weight (kg)	68,00	90,00	75,00	88,00	120,00	165,06	133,03
Family history with overweight (1: yes; 2: no)	1,00	2,00	1,00	1,00	1,00	1,00	1,00
FAVC (1: yes; 2: no)	1,00	2,00	2,00	1,00	1,00	2,00	1,00
FCVC (1: never; 2: sometimes; 3: always)	3,00	3,00	2,00	2,00	2,00	3,00	3,00
NCP (times)	3,00	3,00	3,00	3,00	3,00	3,00	3,00
CAEC (1: no; 2: sometimes; 3: frequenly; 4: always)	3,00	4,00	3,00	4,00	2,00	2,00	2,00
SMOKE (1: yes; 2: no)	1,00	1,00	2,00	1,00	2,00	1,00	1,00
CH2O (liter)	3,00	3,00	2,00	3,00	2,00	2,41	2,98
SCC (1: yes; 2: no)	1,00	1,00	1,00	1,00	1,00	1,00	1,00
FAF (times a week)	1,00	2,00	2,00	3,00	2,00	1,00	1,59
TUE (hours)	1,00	1,00	2,00	1,00	1,00	0,48	0,63
CALC (1: no; 2: sometimes; 3: frequenly; 4: always)	2,00	2,00	2,00	2,00	2,00	3,00	2,00
MTRANS (1: Automobile; 2: Motorbike; 3: Bike; 4: Public Transportation; 5: Walking)	1,00	3,00	2,00	1,00	3,00	3,00	1,00

It can be said that the system has been able to recognize and classify the system even though it is not the best either.

The second testing process uses data testing in the amount of 2041 data. The results of the second testing process are described in Tables 6 and 7. Table 6 describes the confusion matrix of the trial results, while Table 7 is the performance calculation of the second testing process.

Even though it doesn't get very good results, the system can still classify according to clusters. The second testing process obtained a system accuracy value of 74.3%, the precision value was 0.753, while the recall value was 0.743. Misclassification because some data in a cluster have similarities with data in other clusters.

Table 3 Average value of cluster 1–7 for each variable

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Age (year)	20,50	29,50	23,50	32,50	29,78	24,28	26,00
Height (m)	1,70	1,65	1,72	1,67	1,79	1,81	1,66
Weight (kg)	52,00	69,00	81,00	79,00	109,41	136,04	111,93
Family history with overweight (1: yes; 2: no)	2,00	1,00	1,00	1,00	1,00	1,00	1,00
FAVC (1: yes; 2: no)	1,50	1,50	1,00	1,50	1,00	1,00	1,00
FCVC (1: never; 2: sometimes; 3: always)	2,50	3,00	2,00	3,00	2,60	2,50	3,00
NCP (times)	3,50	3,00	2,00	3,00	2,00	3,00	3,00
CAEC (1: no; 2: sometimes; 3: frequenly; 4: always)	2,50	2,00	2,50	2,00	2,00	2,00	2,00
SMOKE (1: yes; 2: no)	1,00	1,00	1,00	1,00	1,00	1,00	1,00
CH2O (liter)	2,00	2,50	2,00	1,00	1,91	2,08	2,77
SCC (1: yes; 2: no)	1,00	1,50	1,00	1,00	1,00	1,00	1,00
FAF (times a week)	1,50	1,50	0,50	1,00	1,97	1,39	0,00
TUE (hours)	1,00	0,50	0,50	0,50	0,00	0,47	0,14
CALC (1: no; 2: sometimes; 3: frequenly; 4: always)	2,00	2,00	2,00	2,50	2,00	2,00	2,00
MTRANS (1: Automobile; 2: Motorbike; 3: Bike; 4: Public Transportation; 5: Walking)	2,00	2,00	1,00	3,00	2,00	1,00	1,00

4 Discussion

The learning and testing process for predicting obesity based on habid data includes frequent consumption of high calorie food (FAVC), frequency of consumption of vegetables (FCVC), Number of main meals (NCP), Consumption of food between meals CAEC, whether smoking or not, Consumption of water daily (CH2O), calories monitoring (SCC), frequency of physical activity (FAF), technological devices used (TUE), frequency of alcohol consumed (CALC), and transportation information used every day have been carried out and testing results have been obtained. Based on the minimum, maximum and average data table (Tables 1, 2 and 3) shows that age allows all ages to be affected by obesity. This is evident that the average age of those affected by obesity is 29.78 for the Type I obesity group, 24.28 years for the Type II obesity group, and 26 years for the Type III obesity group. The age group that

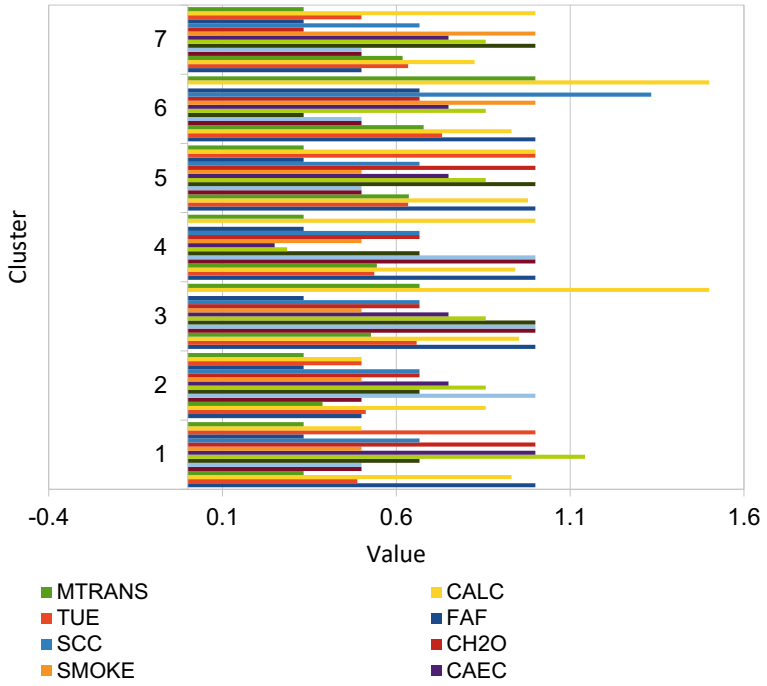


Fig. 1 Illustration of gold standart variable values in clusters 1–7

Table 4 Confusion matrix of the system using learning data

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Cluster 1	8	2	0	0	0	0	0
Cluster 2	0	7	2	1	0	0	0
Cluster 3	0	2	8	0	0	0	0
Cluster 4	0	1	1	9	0	0	0
Cluster 5	0	0	1	1	8	0	0
Cluster 6	0	0	0	0	0	8	2
Cluster 7	0	0	0	0	0	1	9

has a higher prevalence of overweight occurs in the 18+ age group, followed by the 5–12 year age group [27].

Based on the experiments above, it shows that family history factors influence the level of obesity, but this can be anticipated with a healthy lifestyle and other obesity factors must be fully anticipated. This condition is in accordance with research results which show that there is a possibility of obesity in previous family lines, however, the risk of obesity can be overcome by living with a standard lifestyle pattern which must be started from an early age, especially among those who have a family history

Table 5 System performance using learning data

Performance	Cluster	Value
Akurasi	All cluster	0,814
Precision	Cluster 1	1,000
	Cluster 2	0,583
	Cluster 3	0,667
	Cluster 4	0,818
	Cluster 5	1,000
	Cluster 6	0,889
	Cluster 7	0,818
	All cluster	**0,825**
Recall	Cluster 1	0,800
	Cluster 2	0,700
	Cluster 3	0,800
	Cluster 4	0,818
	Cluster 5	0,800
	Cluster 6	0,800
	Cluster 7	0,900
	All cluster	**0,803**

Table 6 Confusion matrix of the system using testing data

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Cluster 1	192	39	29	2	0	0	0
Cluster 2	16	205	48	8	0	0	0
Cluster 3	4	42	200	28	6	0	0
Cluster 4	0	1	37	212	30	0	0
Cluster 5	0	0	1	54	257	29	0
Cluster 6	0	0	2	8	30	212	35
Cluster 7	0	0	0	2	25	48	239

of obesity. This is because genetic predisposition cannot be changed at all over a certain period of time [28].

Meanwhile, the frequency of eating snacks in a day based on experimental results shows that it is a factor that influences the occurrence of obesity. Based on previous research, adults who ate out 1–2 times a day showed a higher risk than those who did not eat snacks at all [29].

Daily physical activity is an important factor that influences obesity levels. This is also in accordance with research results which state the importance of physical activity carried out by a person. Daily activities can be used as initial data to detect the risk of obesity in the future. Therefore, health education regarding the importance

Table 7 System performance using testing data

Performance	Cluster	Value
Akurasi	All cluster	0,743
Precision	Cluster 1	0,906
	Cluster 2	0,714
	Cluster 3	0,631
	Cluster 4	0,675
	Cluster 5	0,739
	Cluster 6	0,734
	Cluster 7	0,872
	All cluster	**0,753**
Recall	Cluster 1	0,733
	Cluster 2	0,740
	Cluster 3	0,714
	Cluster 4	0,757
	Cluster 5	0,754
	Cluster 6	0,739
	Cluster 7	0,761
	All cluster	**0,743**

of daily physical activity by carrying out routine activities such as gymnastics and sports needs to be disseminated to the community [30].

The results of the confusion matrix in the learning process using data learning in Table 4 show that the error in classifying cluster 1 turned out to be classified as cluster 2. This classification error was because some of the data was in the transition area between clusters. This condition occurs in almost all cluster classifications. However, this testing process still has an accuracy of 81.4%, where this condition can be continued for the testing process using testing data.

This phenomenon also occurs in the testing process using testing data. This classification error is due to some data being in the transition area between clusters. This is supported by the facts in the confusion matrix in Table 6. The accuracy of the system is lower than the first testing process, namely 74.3%, the precision value was 0.753, while the recall value was 0.743.

This misclassification occurs because there are indications that some variable values are almost the same in several classes. This is evidenced by the results of the first and second testing confusion matrices, cluster 1 data is classified as cluster 2 and 3, even a little data is classified as cluster 4. This indicates that some data in cluster 1 has similarities, or some variable values overlap with data in cluster 2 and 3. The same thing happened to other clusters.

The limitation and weakness of the results of this experiment is that the input data entered must be complete. Predictions cannot be made if not all variables are

answered completely. However, this research is very useful in that the user can condition the factors that can be anticipated.

5 Conclusion

The aim of this study is proposed obesity prediction approach using a Self Organizing Map by considering habit parameters and clinical variables as input from the system. The learning and testing process has been carried out and the results of the first testing process were 81.4%, the precision value was 0.825, while the recall value was 0.803. The second testing process obtained a system accuracy value of 74.3%, the precision value was 0.753, while the recall value was 0.743. However, this system has been able to classify according to the cluster, even though the accuracy obtained is 74.3%.

Based on the results obtained, it is expected that further research will obtain better accuracy by conducting a sharper analysis of the considered variables. Besides that, it can also be done using machine learning with other methods, even by hybridizing with other methods.

Acknowledgements We are grateful to Universitas Negeri Surabaya for providing the facilities and thanks also to the UCI Machine Learning Repository for supporting the data so that this research can be completed.

References

1. MD KMG, Martin CK, Berthoud H-R, MD SBH (Jan 2018) Obesity: pathophysiology and management. J Am Coll Cardiol 71(1):69–84
2. Romanelli R, Cecchi N, Carbone MG, Dinardo M, Gaudino G, Giudice EMD, Umano GR (Jul 2020) Pediatric obesity: prevention is better than care. Ital J Pediatr 46(103):1–7
3. Sahib AS, Raheem TS, Saleh MS, Khleel LS, Abdulhussein AM, Sadiq SR, Kazaal FAK (2016) Eating behavior in a sample of overweight and obese: a cross sectional study. J Obes Weight-Loss Medicat 2(1):1–6
4. Kinlen D, Cody D, O'Shea D (Jul 2018) Complications of obesity. QJM: Int J Med 111(7):437–443
5. Chaithanya K, Spurthi B, Supreetha A, Yashaswini C (Sep 2020) Obesity, a threat to global health: a review. Int J Res Rev 7(9):265–277
6. Dhurandhar NV (2022) What is obesity? Int J Obes 46:1–2
7. Jiménez EG (Nov 2011)Genes and obesity: a cause and effect relationship. Endocrinología y Nutrición (English Edition) 58(9):492–496
8. Omer T (2020) The health and economic impacts of obesity. Adv Obes Weight Manag Control 10(4):1–3
9. Fruh SM (2017) Risk factors, complications, and strategies for sustainable long-term weight management. J Am Assoc Nurse Pract 29:3–14
10. Jeon J, Lee S, Oh C (Jan 2023) Age-specific risk factors for the prediction of obesity using a machine learning approach. Front Public Health 10(2022):1–12, 17

11. Rodrígueza E, Rodrígueza E, Nascimentoa L, Silvaa AD, Marins F (2021) Machine learning techniques to predict overweight or obesity. In: 4th international conference on informatics & data-driven medicine, Valencia, Spain
12. Mondal PK, Foysal KH, Norman BA, Gittner LS (2023) Predicting childhood obesity based on single and multiple well-child visit data using machine learning classifiers. Sensors (Basel) 23(2):1–27
13. Lim H, Lee H, Kim J (2023) A prediction model for childhood obesity risk using the machine learning method: a panel study on Korean children. Sci Rep 13(10122):1–8
14. UCI Machine Learning Repository (2019) Estimation of obesity levels based on eating habits and physical condition
15. Palechor FM, Manotas ADLH (2019) Dataset for estimation of obesity levels based on eating habits and physical condition in individuals from Colombia, Peru and Mexico. Data Brief 25(104344):1–5
16. Rauber A, Merki D, Dittenbach M (Nov 2002) The growing hierarchical self-organizing map: exploratory analysis of high-dimensional data. IEEE Trans Neural Netw 13(6):1331–1341
17. Dozono H, Niina G (2016) Convolutional self organizing map. In: 2016 international conference on computational science and computational intelligence (CSCI)
18. Kilgore PCSR, Trutschl M, Cvek U, Nam HW (2020)hSOM: visualizing self-organizing maps to accomodate categorical data. In: 2020 24th international conference information visualisation (IV)
19. Aly S, Almotairi S (Jun 2020) Deep convolutional self-organizing map network for robust handwritten digit recognition. IEEE Access 8:107035–107045
20. Suzuki Y, Sakata O, Fukasawa M, Hattori A, Maruyama S, Morisawa M (2018)Auscultating diagnosis support system by using self-organizing map: analysis of long-term recording medical body sounds. In: 2018 joint 10th international conference on soft computing and intelligent systems (SCIS) and 19th international symposium on advanced intelligent systems (ISIS)
21. Delgado S, Morán F, José JCS, Burgos D (2021) Analysis of students' behavior through user clustering in online learning settings, based on self organizing maps neural networks. IEEE Access 132592–132608
22. Tipwong W, Chongjarearn Y, Sirikham A, Preutisrunyanont O (2022) A novel determination of an appropriate clustering quantity of a water-soluble NPK nutrient measuring system based on K-means and SOM methods. In: International electrical engineering congress (iEECON)
23. Kohonen T (1990) The self-organizing map. In: Proceedings of the IEEE
24. Dokmanic I, Parhizkar R, Ranieri J, Vetterli M (Nov 2015) Euclidean distance matrices: essential theory, algorithms, and applications. IEEE Signal Process Mag 32(6):12–30
25. Diego IMD, Redondo AR, Fernández RR, Navarro J, Moguerza JM (Jan 2022) General performance score for classification problems. Appl Intell 52:12049–12063. Springer Link
26. Jiang T, Ping J, Haq AU, Saboor A, Ali A (2021) A novel stacking approach for accurate detection of fake news. IEEE Access 9:22626–22639
27. Oktaviani S, Mizutani1 M, Nishide R, Tanimura S (2021) Prevalence of obesity and overweight stratified by age group of the 34 provinces in Indonesia: local empirical Bayesian estimation. Asian Community Health Nurs Res 3(2):15–21
28. Mangla AG, Dhamija N, Gupta U, Dhall M (Apr 2019) Familial background as a hidden cause for obesity among college going girls. J Biosci Med 7(4):1–13
29. Wei X, Yu D, Ju L, Guo Q, Fang H, Zhao L (Feb 2022) Analysis of the correlation between meal frequency and obesity among Chinese adults aged 18–59 years in 2015. Nutrients 12(3):1–13
30. Suza DE, Miristia V, Hariati H (2020) Physical activities and incidence of obesity among adolescent in Medan, Indonesia. Open Access Maced J Med Sci 20(8):198–203

Entropy-Based Analysis of Electromyography Signal Complexity During Flexion of the Flexor Carpi Radialis Muscle Under Varied Load Conditions



Katherine, Alfian Pramudita Putra, Angeline Shane Kurniawan, Dezy Zahrotul Istiqomah, Nisa'ul Sholihah, Khalid Ali Salem Al-Salehi, Khusnul Ain, Imam Sapuan, and Esti Andarini

Abstract The motor impairments that stroke patients experience as a primary contributor to death and disability also impact their quality of life. Due to its characteristics as complex muscle signals, the Electromyograph (EMG) signal may be one of the markers in rehabilitation and a substitute that may be used to judge the success of rehabilitation. This study aims to analyze the complexity of EMG signals originating from the flexion of the flexor carpi radialis muscle with various weights indicating the rehabilitation state. It is important to monitor the progress of rehabilitation, and it can be done by using complexity analysis. The complexity of the EMG is examined in this work utilizing entropy measurement, namely Approximate Entropy (ApEn), Sample Entropy (SampEn), and Fuzzy Approximate Entropy (fApEn). We concentrated on the flexor carpi radialis muscle using flexion–extension exercises and varying the weight load. This study's results indicate differences in the complexity of the EMG signal in the flexor carpi radialis when performing flexion–extension movements. Three differences showed a correlation based on the value of ApEN, SampEn, and fApEn. Complexity Analysis using fApEn showed the best correlation at R = 0.9966, which can be used to evaluate the rehabilitation process. The increase

Katherine · D. Z. Istiqomah · N. Sholihah · K. A. S. Al-Salehi · K. Ain
Biomedical Engineering Master Program, Department of Physics, Faculty of Science and Technology, Universitas Airlangga, 60115 Surabaya, Indonesia

A. P. Putra (✉) · A. S. Kurniawan · K. Ain
Biomedical Engineering Study program, Department of Physics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia
e-mail: alfian.pramudita@fst.unair.ac.id

I. Sapuan
Department of Physics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia

E. Andarini
School of Nursing, Southern Medical University, Guangzhou 510515, Guangdong, China

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_38

in weight was followed by the increase of the fApEn value, indicating that the value can be used to evaluate the rehabilitation process. The use of complexity analysis of EMG signal from flexion movement in flexor carpi radialis with various weights can be used to indicate the progress in the rehabilitation process.

Keywords Electromyography · Fuzzy approximate entropy · Healthcare · Rehabilitation · Stroke

1 Introduction

Stroke is a disease that affects many people in the community. According to data from the World Stroke Organization's (WSO) Global Stroke Fact Sheet 2022, stroke is the world's second most prevalent cause of death and the third most common cause of disability [1]. Most instances are concentrated in low- and lower-middle-income nations [2]. The estimated global cost of stroke has increased significantly, with a 70% increase in incident strokes, a 43% increase in stroke-related deaths, a 102% increase in general strokes, and a 143% increase in disability-adjusted life years (DALYs) [1]. As a result, stroke cases receive much attention, and treatments are constantly being developed to lessen the impact [3].

As a significant cause of death and disability, stroke causes motor impairments, making it difficult for patients to move [4]. This can substantially impact a person's quality of life, both wisely and financially. Post-stroke rehabilitation is critical in assisting stroke survivors to regain lost or compromised physical functions due to brain damage induced by the stroke [5, 6]. The improvement in motor function in stroke patients after rehabilitation might be considered a measure of rehabilitation success [6, 7]. As a result, throughout patients' rehabilitation, a sensitive and reliable assessment of their motor function is critical [5, 7].

The Manual Muscle Testing (MMT) scoring system is frequently employed in evaluating rehabilitation outcomes, but it is still a manual system [8]. Physiatrists, physiotherapists, or neurologists must measure MMT, and this procedure is unsuccessful and time-consuming. Another way of determining rehabilitation success is ROM measurement, which is a process that still requires expertise and is done manually with a goniometer [9]. As a result, measurement errors frequently result in misdiagnosis, lowering the accuracy of this procedure. So that requires a system that can overcome the scarcity of experts while also conducting an assessment automatically.

The EMG signal is an alternative that can be utilized to assess rehabilitation success [7]. The electrical potential generated during contraction is recorded by EMG [6, 10]. This EMG recording will reveal information on the nervous system's muscle activation; the more muscles contract, the more action potentials are recorded, and the higher the EMG value [11]. As a result, EMG signal analysis is frequently utilized to characterize the electrical activity of muscle fibers during contractions, both isometric and isotonic [12]. EMG signals might be one of the markers in rehabilitation due to their properties as complex muscle signals.

Several studies have been conducted to define the complexity of the EMG signal; one study proposed fuzzy approximate entropy (fApEn) to quantify changes in the complexity of the EMG signal between groups [13]. The results demonstrate that stroke and aging cause neurological alterations. Using the same method in our previous study, we found a substantial difference between the fApEn results on the affected side and the unaffected side using the same way, which refers to the EMG signal's complexity [7]. Meanwhile, the other study compared the FuzzyEn method to other methods, including sample entropy (SampEn). They found that FuzzyEn from EMG showed an increase in a linear correlation with muscle torque rather than Root Mean Square (RMS) EMG amplitude but not with SampEn, which was insensitive to changes in muscle contraction rate [14]. Another study employed a different approach to examine signal complexity in the masseter and temporal muscles, which takes into account the amplitude in the Root Mean Square (RMS), Mean Frequency (MNF), Median Frequency (MDF), and Approximation Entropy (ApEn) [15]. The results reveal a difference in the signal during activity and rest, but there is no linearity in the signal.

According to prior research, entropy measurement is a common way of analyzing the complexity of EMG, particularly the ApEn, fApEn dan SampEn. However, no study utilized EMG signals from muscle contraction under different loading to measure the complexity of the signals. Therefore, we will examine the effectiveness of these three approaches in evaluating the complexity of the EMG signal in this study. To investigate the complexity of the EMG signal, we focused on the upper limb, specifically the flexor carpi radialis muscle. This is because around half of all stroke patients suffer unilateral motor impairments that impair upper-extremity function [5]. As a result, hand and shoulder movements are usually the first to be rehabilitated to support the patient's daily activities. This study aims to differentiate the complexity of EMG signal when someone is doing flexion–extension movements with a correlation of weight load, which is the most common movement done on the upper arm [16, 17]. This study determines the association between EMG signal complexity and weight load in flexor carpi radialis when doing flexion–extension movements. To the extent of our knowledge, there is no study regarding the use of complexity analysis of EMG signal from flexion movement in flexor carpi radialis with various weights to indicate the progress in the rehabilitation process. By having this study, we can contribute to giving more information on the progress during rehabilitation by using complexity analysis.

2 Materials and Methods

In this study, the retrieval and recording of EMG signal data were carried out using Arduino Uno and Olimex shield, which can be described in the block diagram in Fig. 1.

The subject in this study is a normal subject without any muscle impairments, and the subject filled out the informed consent before the experiment was conducted.

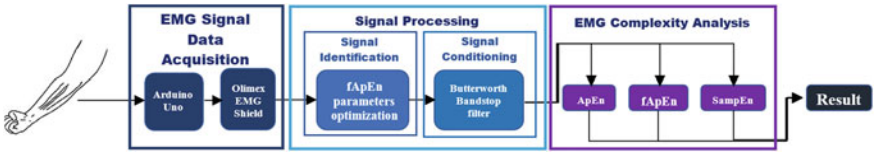


Fig. 1 Block diagram of EMG signal acquisition and analysis

The subject flexed their wrist in 10 repetitions, and their EMG signal from the flexor carpi radialis muscle was recorded. The data acquisition process was carried out by placing a disposable bipolar Ag/AgCl electrode (Onemed, Indonesia) on the flexor carpi radialis muscle, with the position of the electrode placement based on SENIAM recommendation. The active electrode was placed on the line between the medial acromion and the fossa cubit at 1/3 from the fossa cubit. The reference electrode was on the tendon position near the wrist, and the ground electrode was near the elbow.

After the electrode was placed, the hand was moved until it approximately reached an angle of 90° and back to its original position. This experiment was performed with three load variations, which were no load, 1, and 2 kg. The flexor carpi radialis muscle was in contraction and relaxation condition by flexing the hand to see the difference in EMG signal based on its amplitude. The duration for flexion and extension was around 5 s for each section.

In this study, we used Fuzzy Approximate Entropy (fApEn) to evaluate the complexity of the EMG signal. We compared them with the other two entropy-based complexity analyses, which were Approximate Entropy (ApEn) and Sample Entropy (SampEn). According to Cirugeda-Roldan et al. [18], the weakness of ApEn is that it calculates self-match and depends on the signal length [18]. However, ApEn can provide clinically significant information from biomedical records. According to previous studies, ApEn can be translated into the form of Eqs. 1–8 [19–21].

A vector sequence X_i^m with length of m is obtained from T containing data of an N number with a time series {u(i): $1 \leq i \leq N$}.

$$X_i^m = \{u(i), u(i+1), \ldots, u(i+m-1)\}, 1 \leq i \leq N-m+1 \tag{1}$$

where m successive u values starting with the i-th point starting X_i^m. Equation 2 is the definition of the distance d_{ij}^m between X_i^m and X_j^m.

$$d_{ij}^m d[X_i^m, X_j^m] = \underset{k \in (0, m-1)}{maks} |u(i+k) - u(j+k)| \tag{2}$$

Equation 3 is a measure that defines the similarity C_r^m between vector X_i^m and vector X_j^m.

$$C_r^m = \frac{1}{N-m+1} \sum_{j=1, j\neq i}^{N-m+1} \Theta(d_{ij}^m - r) \tag{3}$$

where Heaviside functions as Θ.

$$\Theta(z) = \begin{cases} 1 \, if \, z \leq 0 \\ 0 \, if \, z > 0 \end{cases} \tag{4}$$

In Eq. 3, there is a symbol r representing a predetermined tolerance value.

$$r = k.std(T) \tag{5}$$

The constant is denoted k (k > 0), and the time series standard deviation is as std Θ. The value of r will be plugged into Eq. 6 to be defined as ϕ^m.

$$\phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln[C_r^m(i)] \tag{6}$$

Based on the time series, the ApEn value can be calculated using Eq. 7.

$$ApEn(m, r) = \lim_{N\to\infty} [\phi^m(r) - \phi^{m+1}(r)] \tag{7}$$

The time series ApEn value can be obtained for practical applications using N data such as Eq. 8.

$$ApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r) \tag{8}$$

In 2002, Richman and Moorman proposed the first SampEn plan as a solution to reduce bias on ApEn. On random number theory for various operating conditions, SampEn is better than ApEn. In addition, the SampEn calculation algorithm is more straightforward than ApEn [18]. SampEn differs from ApEn in two ways: by not calculating selfby-matches and not using a template-wise approach [22].

SampEn's calculation algorithm is more straightforward than ApEn's. The first and second steps are the same, but Eq. 9 becomes:

$$C_r^m = \frac{1}{N-m-1} \sum_{j=1, j\neq i}^{N-m} \Theta(d_{ij}^m - r) \tag{9}$$

$$\phi^m(r) = \frac{1}{N-m-1} \sum_{i=1}^{N-m} C_r^m(i) \tag{10}$$

Thus, SampEn is defined as:

$$SampEn(m, r, N) = ln\left(\phi^m(r)\right) - ln(\phi^{m+1}(r)) \tag{11}$$

The fApEn method was proposed to characterize time series regularities and overcome poor statistical stability in ApEn and SampEn [7, 18]. Zadeh introduces the concept of a "fuzzy set", which proposes a way to characterize the input–output relationship in an imprecise environment [7, 19]. Also, a membership function is introduced by fApEn, and then a pattern belongs to class Z when evaluated by $\mu_Z(x)$. The higher the membership value of the x pattern belonging to a class, the closer the matter is to 1 [18]. The fApEn method is defined as:

$$\begin{aligned} X_i^m &= \{u(i), u(i+1), \ldots, u(i+m-1)\} - u0(i) \\ &(i = 1, \ldots, N-m+1) \end{aligned} \tag{12}$$

Then the baseline is deleted to generalize X_i^m

$$u0(i) = \frac{1}{m}\sum_{j=0}^{m-1} u(i+j) \tag{13}$$

Then, Eq. 14 is the distance between two vectors X_i^m and X_j^m.

$$\begin{aligned} d_{ij}^m &= d\left[X_i^m, X_j^m\right] = \max_{k\epsilon(0,m-1)} |u(i+k) - u0(i) - (u(j+k) - u0(j))| \\ &(i, j = 1 \sim N-m+1, j \neq i) \end{aligned} \tag{14}$$

The fuzzy membership function determines the degree of similarity between D_{ij}^m between X_i^m and X_j^m.

$$D_{ij}^m = u\left(d_{ij}^m, r\right) \tag{15}$$

On the B_r^m function is determined by

$$C_r^m(i) = \frac{1}{N-m+1} \cdot \sum_{j=1, j\neq i}^{N-m+1} D_{ij}^m \tag{16}$$

$$\phi^m(r) = \frac{1}{N-m+1} \cdot \sum_{i=1}^{N-m+1} \ln\left[C_r^m(i)\right] \tag{17}$$

To define fApEn (m, r) based on a time series, a vector series $\{X_i^{m+1}\}$ and a function $\phi^{m+1}(r)$ are used.

$$fApEn(m, r) = \lim_{N\to\infty}\left[\phi^m(r) - \phi^{m+1}(r)\right] \tag{18}$$

Thus, statistically fApEn can be defined like Eq. 18 for a finite dataset.

$$fApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r) \tag{19}$$

We use some simulated signals to test the fApEn algorithm before being used in EMG signals. The simulated signals were called MIX (p) shown in Eq. 20.

$$MIX\,(p)_n = (1 - A_n)B_n + A_nC_n \qquad (1 \leq n \leq x) \tag{20}$$

where A_n is an independent identically distributed real random variable $B_n = \sqrt{2}\sin\left(\frac{2\pi n}{12}\right)$, and C_n is a set of independent identically distributed real random variables with a uniform distribution inside $\left[-\sqrt{3}, \sqrt{3}\right]$. $A_n = 1$ has a probability of p and, $A_n = 0$ has a probability of $1 - p$. When the probability (p) increases, the randomness or complexity also increases in the $MIX\,(p)$. The signals were shown in Fig. 2.

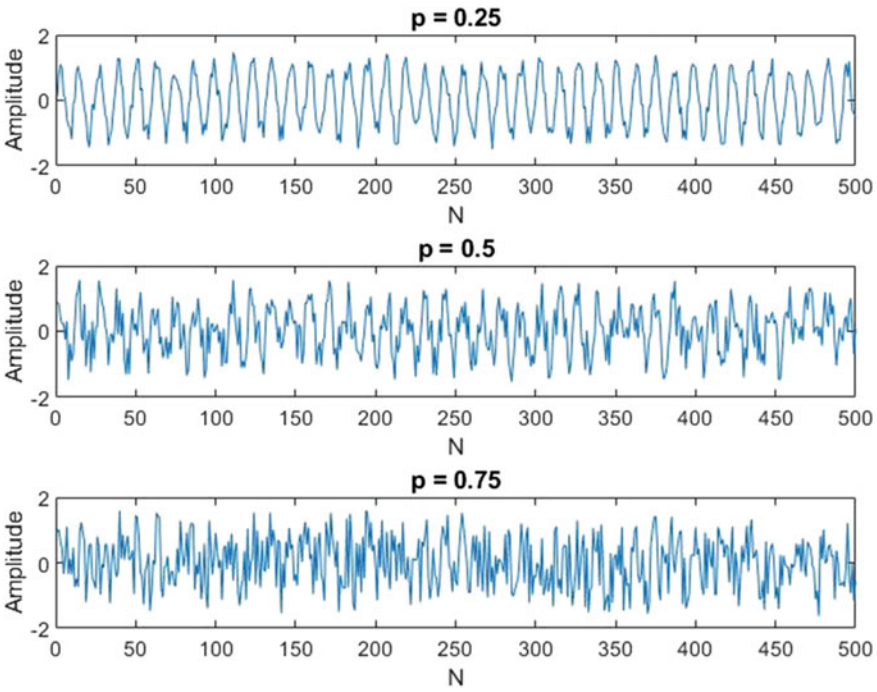


Fig. 2 Simulated signal with different probability which showed different complexity

3 Results

Before measuring the complexity of the EMG signals, the fApEn parameters were optimised by using the simulated signals mentioned before, which already had different stated complexity based on probability. The fApEn method's result will be relative continuity, monotonicity, and consistency. Parameters n and r in the fApEn process were optimised, which will be used later in the EMG signals. The result is shown in Figs. 3 and 4.

The data acquisition results from EMG signals still contain a lot of noise. This condition was caused by the electrode's function as sensors of various forms of artefacts such as baseline wander, noise due to random body movements or movements of the electrodes themselves, and PLI noise [23]. The filtering step using a Butterworth bandstop filter is applied to remove noise from the EMG signal. The filtered EMG signals show different signal morphology, which is illustrated in Fig. 5. It can be seen that the heavier the weight carried, the higher the EMG signal generated and the more comprehensive the contraction range. Muscle tension data was obtained by processing the filtered EMG signal data using fApEn, ApEn, and SampEn. The processing results for each load were shown in Fig. 6.

The average results of fApEn, ApEn, and SampEn processing showed a correlation with the weight of the load. As shown in Fig. 7, it can be seen that as the

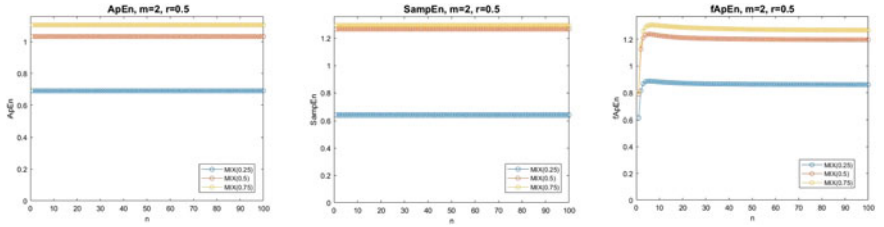


Fig. 3 ApEn, SampEn, and fApEn of three simulated signal MIX (p) with a variation of n

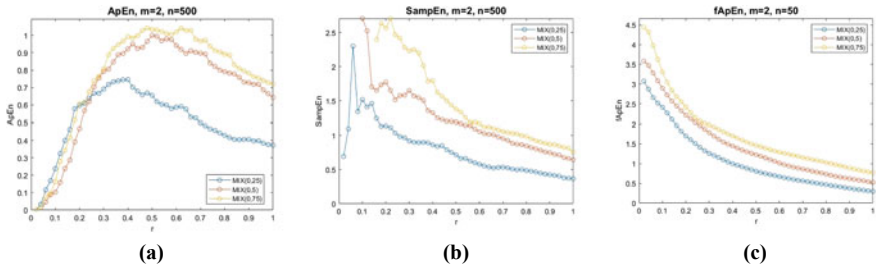


Fig. 4 ApEn, SampEn, and fApEn of three simulated signal MIX (p) with a variation of r

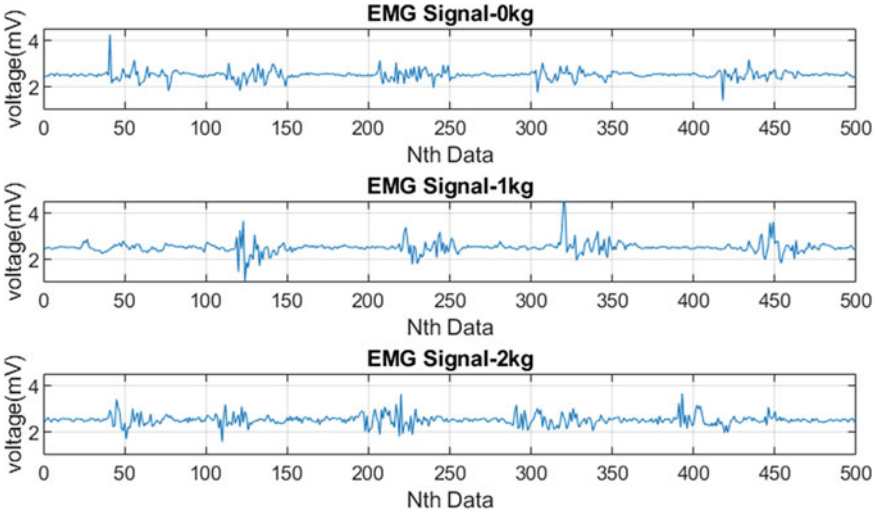
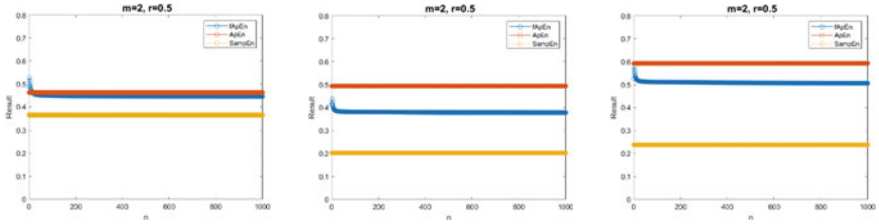


Fig. 5 EMG signal from extension-flexion movement with weight 0, 1, and 2 kg



(a) **(b)** **(c)**

Fig. 6 Result of fApEn, ApEn, and SampEn from EMG signal with weight **a** 0 kg, **b** 1 kg, and **c** 2 kg

weight of the load increases, the values of fApEn, ApEn, and SampEn also consistently increase. However, the difference in the resulting values is not significant. The linearity calculation of R shows fApEn as the best method in recognizing the complexity of EMG signal with an r value reaching 0.9966, followed by SampEn at 0.9592, and the lowest is ApEn at 0.6775.

4 Discussions

In this study, ApEN, SampEn, and fApEn were compared in terms of complexity analysis in EMG signals. fApEn in EMG signal with load 2 kg was the highest compared to the other two conditions. Based on the test using simulated signals,

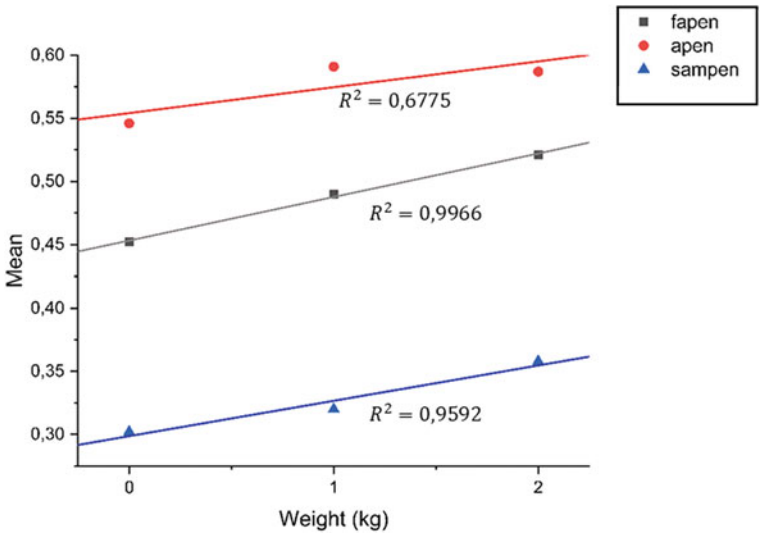


Fig. 7 Comparison of Mean Linearity of fApEn, ApEn, and SampEn

fApEn was superior compared to the ApEn and SampEn, especially when a different r variable was used. The use of fuzzy function in the fApEn can eliminate the drawback that occurred in the ApEn and SampEn which used the Heaviside function [24]. Heaviside function has a rigid limit in the function and it only can be used in smaller data. On the other hand, the fuzzy function has a flexible limit and can be used in longer data so that it will be more convenient for the signal analysis process [25]. The result in this study were aligned with the result in the previous studies which used the simulated signals with assigned complexity by using probability. The result depicted that the signals with a high p-value will have a high fApEn value, indicating high complexity [7, 13, 20, 21]. On the other hand, a signal with a low p-value showed a lower fApEn value. There was an abrupt change in the fApEn value if we used an n value between 1 and 5 based on Fig. 4. After that, the fApEn value remained steady until a high number of n. Besides, when we tried several values of r, the fApEn value decreased rapidly with r less than 0.1.

The complexity of EMG signals from flexor carpi radialis muscle with three load conditions showed that there was a positive linear trend. This result was in line with the other study which also demonstrated the increase in complexity when contraction of the muscle increased [26]. More muscle contraction indicates more firing rates in the motor units. The result also happened if the motor units suffer from pathological condition, such as in post-stroke patients. The fApEn value in post-stroke patients EMG signals was lower than EMG signals in healthy subject [13].

Complexity information can be used as a tool to quantify the EMG signal. The use of different weights was used to indicate the increasing condition of EMG signals. This condition can depict the progress in rehabilitation process. Rulaningtyas et al. study also showed the same behavior as this study. They showed that the complexity

of EMG signals in stroke patient was higher in the unaffected side than that of the affected side [7]. That result also indicated that the complexity of EMG signals can distinguish two level of conditions, even when it was compared to normal condition. By having a good linearity, the result of this study showed that complexity can showed stage of contraction triggered by various load. The higher the load, the more EMG signals will be needed. In healthy human being, a higher complexity should be obtained. The lower entropy indicated the impairment condition of a person [12]. Another study conducted by Ao et al. showed that the change in complexity can be detected in different groups based on stroke and aging condition [13].

The result showed a positive trend that complexity analysis can be used to quantify the EMG signals so that it can be used to evaluate the rehabilitation progress. However, some limitations need to be taken into account. The complexity analysis in this study was performed offline which the EMG data was analyzed at a different time so that it is still not real-time. Some improvement needs to be taken to make the process shorter and it can be used in real-time condition. For future study, the complexity of EMG signals based in various weight can be used as input for actuating orthosis or prosthetics.

5 Conclusions

This research shows differences in the complexity of the EMG signal when a person performs a flexion–extension movement that is correlated with a heavy load, which is indicated by the results of EMG signal processing using fApEn, ApEn, and SampEn. Where the relationship between the complexity of the EMG signal and the heavy load is directly proportional and is evidenced by the linearity value of r, where fApEn is the best method for recognizing the complexity of the EMG signal with an R-value of 0.9966, followed by SampEn and ApEn. Variations in weight and muscle movement need to be re-examined to learn more about the complexity of the EMG signal.

For future study, there is a need to evaluate the complexity of EMG signals in several conditions, such as some movement in the activity of daily living, specific movement related to a pathological condition, and movement for a specific purpose. The use of complexity analysis also can be used to determine the status of rehabilitation progress in the patient.

References

1. Feigi VL, Brainin M, Norrving B, Martins S, Sacco RL, Hacke W, Fisher M, Pandian J, Lindsay P (2022) World Stroke Organization (WSO): global stroke fact sheet 2022. Int J Stroke: Off J Int Stroke Soc 17(1):18–29
2. Zheng S, Yao B (2019) Impact of risk factors for recurrence after the first ischemic stroke in adults: a systematic review and meta-analysis. J Clin Neurosci 60:24–30

3. Moussaddy A, Demchuk AM, Hill MD (2018) Thrombolytic therapies for ischemic stroke: triumphs and future challenges. Neuropharmacology 134:272–279
4. Mane R, Chouhan T, Guan C (2020) BCI for stroke rehabilitation: motor and beyond. Neural Eng 17(4)
5. Rahmatillah A, Rahma ON, Amin M, Wicaksana SI, Ain K, Rulaningtyas R (2018) Post-stroke rehabilitation exoskeleton movement control using EMG signal. Int J Adv Sci Eng Inf Technol 8(2):616–621
6. Sun R, Song R, Tong KY (2014) Complexity analysis of EMG signals for patients after stroke during robot-aided rehabilitation training using fuzzy approximate entropy. IEEE Trans Neural Syst Rehabil Eng: Publ IEEE Eng Med Biol Soc 22(5):1013–1019
7. Rulaningtyas R, Prasetyo AB, Rahmatillah A, Putra AP, Rahma OR, Ain K, Pawana IPA (2021) Measuring the complexity of EMG signal by using fuzzy approximate entropy in post-stroke patients rehabilitation. J Eng Sci Technol Rev 14(3):85–90
8. Roman NA, Miclaus RS, Nicolau C, Sechel G (2022) Customized manual muscle testing for post-stroke upper extremity assessment. Brain Sci 12:457
9. Pediyanti MC, Rulaningtyas R, Rahmatillah A, Katherine (2021) Range of motion measurement of Articulatio cubiti based on Hough transformation. In: International conference on mathematics, computational sciences and statistics 2020. AIP conference proceedings, vol 2329
10. Elamvazuthi EI, Duy NHX, Ali Z, Su SW, Khan MKAA, Parasuraman S (2015) Electromyography (EMG) based classification of neuromuscular disorders using multi-layer perceptron. Procedia Comput Sci 76:223–228
11. Enoka RM (2017) Electromyography (EMG). Reference module in neuroscience and biobehavioral psychology
12. Rampichini S, Vieira TM, Castiglioni P, Merati G (2020) Complexity analysis of surface electromyography for assessing the myoelectric manifestation of muscle fatigue: a review. Entropy 22:529
13. Ao D, Sun R, Tong K-Y, Song R (2015) Characterization of stroke- and aging-related changes in the complexity of EMG signals during tracking tasks. Ann Biomed Eng 43:990–1002
14. Zhu X, Zhang X, Tang X, Gao X, Chen X (2017) Re-evaluating electromyogram–force relation in healthy biceps brachii muscles using complexity measures. Entropy 19
15. Giannasi LC, Politti F, Dutra MTS, Tenguan VLS, Silva GRC, Mancilha GP, da Silva DB, Oliveira LVF, Oliveira CS, Amorim JBO, Salgado MAC, Gomes MF (2020) Intra-Day and Inter-Day reliability of measurements of the electromyographic signal on masseter and temporal muscles in patients with Down syndrome. Sci Rep 10:7477
16. Daud WMBW, Yahya AB, Horng CS, Sulaima MF, Sudirman R (2013) Features extraction of electromyography signals in time domain on biceps brachii muscle. Int J Model Optim 515–519
17. Chaytor CP, Forman D, Byrne J, Loucks-Atkinson A, Power KE (2020) Changes in muscle activity during the flexion and extension phases of arm cycling as an effect of power output are muscle-specific. PeerJ 8:e9759
18. Cirugeda-Roldan E, Cuesta-Frau D, Miro-Martinez P, Oltra-Crespo S (2014) Comparative study of entropy sensitivity to missing biosignal data. Entropy 16:5901–5918
19. Xie HB, Guo JY, Zheng YP (2010) Fuzzy approximate entropy analysis of chaotic and natural complex systems: detecting muscle fatigue using electromyography signals. Ann Biomed Eng 38(4):1483–1496
20. Chen W, Wang Z, Ren X (2006) Characterization of surface EMG signals using improved approximate entropy. J Zhejiang Univ Sci B 7(10):844–848
21. Chen W, Zhuang J, Yu W, Wang Z (2009) Measuring complexity using FuzzyEn, ApEn, and SampEn. Med Eng Phys 31(1):61–68
22. Chen W, Wang Z, Xie H, Yu W (2007) Characterization of surface EMG signal based on fuzzy entropy. IEEE Trans Neural Syst Rehabil Eng 15:266–272
23. De Luca CJ, Donald Gilmore L, Kuznetsov M, Roy SH (2010) Filtering the surface EMG signal: movement artifact and baseline noise contamination. J Biomech 43(8):1573–1579

24. Zhu P, Wu Y, Liang J, Ye Y, Liu H, Yan T, Song R (2018) Characterization of the stroke-induced changes in the variability and complexity of handgrip force. Entropy 20
25. Arjunan SP, Kumar DK (2018) Measuring complexity in different muscles during sustained contraction using fractal properties of SEMG signal. In: Annual international conference of the IEEE engineering in medicine and biology society, pp 5656–5659
26. Chen Y, Hu H, Ma C, Zhan Y, Chen N, Li L, Song R (2018) Stroke-related changes in the complexity of muscle activation during obstacle crossing using fuzzy approximate entropy analysis. Front Neurol 9:131

Voice Features Examination for Parkinson's Disease Detection Utilizing Machine Learning Methods



Farika Tono Putri, Muhlasah Novitasari Mara, Rifky Ismail, Mochammad Ariyanto, Hartanto Prawibowo, Triwiyanto, Sari Luthfiyah, and Wahyu Caesarendra

Abstract Manageable symptoms can be a critical and important thing for people with Parkinson's disease (PWP) in order to maintaining their quality of life. PWP early detection and monitoring is one of way to understand whether the medication dosage and physical therapy manage to maintain the stage of symptoms. The cheapest monitoring method can be done is based on voice signal. PD detection method based on voice signal shows promising future to be implemented into real world through an online and mobile based medical related application. However, It is necessary to select the important voice features which contribute to highest detection accuracy so that the implementation of detection through the application can be done effectively without requiring complex mathematical code. This study aim to evaluate and analize the highest accuracy and suitable voice feature related to PWP early detection and monitoring using machine learning method. Voice data recorded from study participants consist of Hughes-based stages of Parkinson's disease (PD)

F. T. Putri · M. N. Mara
Politeknik Negeri Semarang, Semarang, Indonesia

F. T. Putri · R. Ismail · M. Ariyanto · H. Prawibowo
Center for Bio-Mechanics Bio-Material Bio-Mechatronics and Bio-Signal Processing (CBIOM3S), Semarang, Indonesia

R. Ismail · M. Ariyanto · H. Prawibowo
Universitas Diponegoro, Semarang, Indonesia

M. Ariyanto
Osaka University, Suita, Japan

Triwiyanto · S. Luthfiyah
Department of Electromedical Engineering, Poltekkes Kemenkes Surabaya, Surabaya, Indonesia

W. Caesarendra (✉)
Faculty of Integrated Technologies, Universiti Brunei Darussalam, Jalan Tungku Link, Gadong, 1410 Brunei Darussalam, Brunei
e-mail: wahyu.caesarendra@ubd.edu.bn

Faculty of Mechanical Engineering, Opole University of Technology, 76 Proszkowska St., 45-758 Opole, Poland

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_39

patients and healthy subjects. Participants recorded their voice said 'aaaa..' for 5 s then the voice data calculated into 22 voice features. Those features then examine using machine learning methods such as logistic regression, random forest, KNN and deep learning CNN and classified into four classes based on Hughes standard e.g. healthy, possible, probable and definite. The experimental result showed that the most suitable features are 11 features out of 22 features which examined using random forest and CNN method contributed to highest accuracy value of 95%. This most important features then can be implemented along with CNN method into an online and mobile based application for future study.

Keywords Parkinson's disease detection · Voice feature · Machine learning

1 Introduction

Parkinson's Disease (PD) is a neurodegenerative disease with tremor, rigidity, bradykinesia and postural instability symptoms [1]. Those symptoms often accompanied by another problems related to both physical and mental aspect. Depression, emotional changes and sleep problems are the problem related to mental aspect of people with PD (PWP). Meanwhile, some problems related to physical aspects are swallowing and chewing difficulties, changes of speech, urinary problems and constipation as well as skin problems. Some of the PD's accompanied problems can be managed and treated using drugs and physical therapy [1, 2]. Therefore, the PD early detection and symptoms monitoring can be a crucial method in order to manage PWP symptoms and their respond to drugs and physical therapy. However, PD detection and monitoring in Indonesia mostly rely on neurologist clinical examination accompanied by some motor impairment symptoms presence such as tremor, rigidity, bradikinesia and postural instability and non-motor impairment [3]. Those clinical examination was subjective and the examination result can be different between one neurologist to another depend on the scale used [4]. Another method to detect and monitor PD stages can be conducted using electromyography (EMG) signals [5, 6] or another image based detection such as positron emission tomography (PET) scan and magnetic resonance imaging (MRI) [7, 8]. However, EMG and image based PD detection and monitoring can be costly to patients. One of PD monitoring method which low cost and easy to use can be conducted with voice based detection using machine learning [3–5]. Voice based PD detection and monitoring can be an innovative method to be implemented into telehealth mobile based application. In order to design a fast algorithm and effective telehealth application, PD detection code need to be simple yet accurate. Previous study related to this research has been conducted to build PD early detection and monitoring based on voice signal using 5 voice features with Artificial Neural Networks (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) as classification method [9]. The overall accuracy for PD four classes based on Hughes scale resulted at 81.75% on ANN method and 90.48% on NFIS method [9]. This accuracy result can be improved with features addition

and examination using various machine learning methods in purpose to telehealth application in the future.

Previous studies related to early PD detection and monitoring based on voice signals have been conducted with various and extensive methods. Most studies related to PD detection and monitoring classified participants into two classes, i.e. healthy and PD [7, 10–22]. Meanwhile there was not many studies classified PD detection and monitoring into multiclass. Multiclass PD detection and study mostly based on The Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS) with three classes category [23] (healthy-mild based on MDS-UPDRS III-severe based on MDS-UPDRS S) and five class [24]. Another study [25] classify two class (early stage without L-Dopa and mid stage with L-Dopa) based on Hoehn and Yahr scale. These literatures set differences with this study which classify PD detection into four classes (healthy, possible, probable and definite) based on Hughes scale. Hughes scale mostly use in Indonesia as clinical examination guidance for neurologist.

Various machine learning methods and accuracy results have been proposed and obtained from previous study. Table 1 shows the summary of various PD detection and monitoring machine learning method, number of features used, dataset, number of classes and their accuracy.

Table 1 shows that more than one recent study used UCI repository PD dataset. Another studies using the same dataset with various machine learning method obtained 97.5% accuracy using variable adaptive moment-based backpropagation algorithm of ANN called BPVAM [30]. Another study [31] obtained 92.19 and 93.84% accuracy for DNN and FESA-DNN method.

This presented paper fill in the gap where not many previous study conducted which is multiclass classification for PD detection and monitoring. This study also using Hughes scale based which not used a lot in another country outside of Indonesia. This study used Indonesian participants. For feature selection this paper used importance value which assign a score to input features based on how useful they are at predicting a target variable. The aim of this study is to examine voice feature effect on algorithm accuracy for PD stages detection. This study contribute to voice based PD detection for Hughes scale (possible-probable-definite-healthy) with best machine learning method and voice features combination.

The result of this research can be a first step into an effective and efficient telehealth application development.

2 Methodology

Research methodology for this study consist of: (i) Voice data acquisition, (ii) Feature calculation and (iii) Voice feature examination using machine learning methods.

Voice data acquisition recorded from 23 research participants whom 15 participants are PWP patients and 8 participants are healthy participants as presented in Table 2. Research participants consist of both male and female gender with age

Table 1 PD detection and monitoring literatures summary

Year published	Machine learning method	Number of voice features	Dataset	Number of classes	Accuracy (%)
2023 [10]	IBk	N/A	University of Oxford [26]	2	96.41
	Naïve Bayes				69.23
	Multilayer Perceptron (MLP)				90.77
	Random Forest (RF)				91.79
	Random Tree				84.62
	Sequential Minimal Optimization (SMO)				87.18
	Smallest Normalized Difference Associative Memory (SNDAM)				96.61
	Improved SNDAM (ISNDAM)				99.48
2023 [10]	IBk	N/A	Istanbul University [27]	2	67.96
	Naïve Bayes				59.85
	MLP				68.05
	RF				73.10
	Tandom Tree				66.14
	SMO				66.23
	SNDAM				98.28
	ISNDAM				99.66
2023 [12]	End to end deep learning	22	University of California, Irvine (UCI) Repository [28]	2	90
2023 [23]	Support Vector Machine (SVM)	4	Spanish PD database PC-GITA [29]	3	45
	RF				48
	MLP				48
	AdaBoost				47

(continued)

Table 1 (continued)

Year published	Machine learning method	Number of voice features	Dataset	Number of classes	Accuracy (%)
2023 [11]	K-Nearest Neighbour (KNN) with gridsearch and feature selection	22	UCI Repository [28]	2	88
	SVM with gridsearch				83
	Decision Tree with gridsearch				90
	RF with gridsearch				92
	MLP with gridsearch				98

range from 39 to 80 years old. The youngest PWP patient with early onset symptoms is a 39 years old woman. PWP data were collected from Dr. Kariadi General Hospital, Semarang, Indonesia. PD status distinguished based on Hughes criteria where possible status determined when patient shows one of main symptom such as rigidity, bradukinesia, tremor at rest or postural reflexes failure [32]. Probable stage determine when patient suffers a combination of two main symptoms including postural reflexes failure or asymmetric rigidity or bradykinesia [32]. Definite stage when patient shows combination of three out of four main symptoms or combination of two main symptoms with one symmetric symptom [32].

PWP suffers voice dysfunction as an earliest sign of motor impairment. Those voice dysfunction is a phenomena associated with rigidity and bradykinesia which happened in laryngeal muscles [33]. The voice dysfunction which some called hypokinetic dysarthia have some symptoms that affected PWP, i.e. reduced volume, hoarse voice and tremor phonation [34]. Based on reason mentioned, voice data acquisition conducted as follows:

1. Research participants produce phonation "aaaa….." with stable pitch within 5 s.
2. Data recording was done in six times for each research participants.

Figure 1 shows the voice data acquisition process between healthy participants (Fig. 1a) and PWP participants (Fig. 1b) [9].

Unidirectional microphone is used for voice data recording based on consideration that the microphone is sensitive to voice with only one direction so it can reduce another noise from environment. This fact is backed with previous research which used unidirectional microphone in purpose for respiration sound recording [35]. Voice recording conducted with 44 kHz sampling rate.

After voice data recording, voice feature calculation was conducted. There were 22 voice features used in this study. All those voice features are the baseline features showed in Table 3 [17].

Table 2 Research participants' data

PD participants				Healthy participants		
Subject	Gender	Age (y.o)	PD status	Subject	Gender	Age (y.o)
S1	Male	68	Definite	S16	Female	71
S2	Male	53	Possible	S17	Female	53
S3	Male	59	Possible	S18	Female	75
S4	Female	79	Definite	S19	Male	70
S5	Female	66	Definite	S20	Female	55
S6	Female	39	Probable	S21	Female	70
S7	Female	58	Possible	S22	Male	52
S8	Female	54	Probable	S23	Male	60
S9	Male	66	Probable			
S10	Male	80	Probable			
S11	Male	76	Probable			
S12	Male	70	Definite			
S13	Male	68	Probable			
S14	Male	68	Definite			
S15	Male	72	Definite			



(a) **(b)**

Fig. 1 Voice data acquisition: **a** Healthy participants and **b** PWP participants [9]

Table 3 Voice features review [17]

Measure	Explanation	Feature names
Jitter related parameters	Capturing oscilating pattern irregularities of vocal fold and quantify cycle to cycle changes	Jitter local
		Jitter absolute Relative average perturbation (RAP) Period perturbation quotient 5 (PPQ 5)
Shimmer related parameters	Capturing oscilating pattern irregularities of vocal fold and quantify cycle to cycle changes	Shimmer
		Shimmer (dB) Amplitude perturbation quotient 3 (APQ3) APQ5 APQ11
Fundamental frequency related parameters	Capturing the vibration frequency of vocal folds	Multi-dimensional voice program (MDVP) F0
		MDVP Fhi
		MDVP Flo
		Spread1
		Spread2
		Correlation dimensions
Harmonicity related features	Capturing the noise components due to pathologies related with voice	Noise to harmonic ratio (NHR)
		Harmonic to noise ratio (HNR)
Recurrence period density entrophy (RPDE)	Capturing information about vocal cord oscillation ability	RPDE
Detrended fluctuation analysis (DFA)	Capturing information related to turbulent noise	DFA
Pitch period entrophy (PPE)	Capturing information related to impaired fundamental frequency based on logarithmic scale	PPE
Amplitude and pitch parameters	Capturing irregularities on amplitude and pitch of fundamental frequency	
		Difference of difference of amplitude (DDA)
		Difference of difference of period (DDP)

Jitter related parameter features as shown in Table 3, i.e. jitter local and jitter absolute can be represented using Eqs. (1) and (2) [36].

$$\text{Jitter local} = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|T_i - T_{i-1}|}{\frac{1}{N}\sum_{i=1}^{N} T_i} \times 100 \tag{1}$$

where T_i represented the extracted glottal period lengths and N is number of extracted glottal period.

$$\text{Jitter absolute} = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|T_i - T_{i-1}|}{\frac{1}{N}\sum_{i=1}^{N} T_i} \tag{2}$$

Meanwhile RAP feature defined as difference between period and the average of it and its two neighbor in absolute form, then divided by the average period as seen in Eq. (3) [36]. RAP feature stated in percentage form.

$$\text{RAP} = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}\left|T_i - \left(\frac{1}{3}\sum_{n=i-1}^{i+1} T_n\right)\right|}{\frac{1}{N}\sum_{i=1}^{N} T_i} \times 100 \tag{3}$$

PPQ5 feature is difference between period and the average of it and its four neighbor in absolute form, then divided by the average period. PPQ5 feature represented in Eq. (4) [36]. PPQ5 feature also stated in percentage form.

$$\text{PPQ5} = \frac{\frac{1}{N-1}\sum_{i=2}^{N-2}\left|T_i - \left(\frac{1}{5}\sum_{n=i-2}^{i+2} T_n\right)\right|}{\frac{1}{N}\sum_{i=1}^{N} T_i} \times 100 \tag{4}$$

Shimmer related feature is variation of consecutive period amplitude. There are five measurement related to shimmer: Shimmer local, Shimmer (dB), APQ3, APQ5 and APQ11. Shimmer local define as difference between consecutive periods amplitude in absolute form, divided by average amplitude. Shimmer local denoted in Eq. (5) [36].

$$\text{Shimmer local} = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|A_i - A_{i+1}|}{\frac{1}{N}\sum_{i=1}^{N} A_i} \times 100 \tag{5}$$

where A_i is amplitude and N is fundamental frequency periods.

Shimmer (dB) feature is difference between consecutive period amplitude in 10 based logarithmic form multiplied by 20 as seen in Eq. (6) [36].

$$\text{Shimmer (dB)} = \frac{1}{N-1}\sum_{i=1}^{N-1}|20 \times \log(A_{i+1}/A_i)| \tag{6}$$

APQ3 feature described as difference between a period amplitude and it neighbours amplitude average, divided by average amplitude as denoted in Eq. (7) [36].

$$\text{AQP3} = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}\left|A_i - \left(\frac{1}{3}\sum_{n=i-1}^{i+1} A_n\right)\right|}{\frac{1}{N}\sum_{i=1}^{N} A_i} \tag{7}$$

APQ5 feature defined as difference between a period amplitude and its four closest neighbor, divided by the average amplitude [36]. PPQ5 can be calculated using Eq. (8).

$$\text{APQ5} = \frac{\frac{1}{N-1}\sum_{i=2}^{N-2}\left|A_i - \left(\frac{1}{5}\sum_{n=i-2}^{i+2} A_n\right)\right|}{\frac{1}{N}\sum_{i=1}^{N} A_i} \times 100 \tag{8}$$

Next step after the voice features calculation was voice feature examination using machine learning methods. There were four machine learning methods used in this study, i.e. logistic regression, random forest, KNN and CNN. The machine learning methods was used to classified data into four class based on Hughes scale which mostly used in Indonesia, i.e. possible, probable, definite and healthy.

This study used multiclass logistic regression which defined in Eq. (9) [37] where $y_{nk} = y_n\,(\varnothing_n)$ is softmax function for binary logistic regression as stated in Eq. (10).

$$P(\text{T}|\text{w1},\ldots\text{wK}) = \prod_{k=1}^{N} \text{p(Ck}|\text{Øn)tnk} = \prod_{n=1}^{N}\prod_{k=1}^{K} y_{nk}^{tnk} \tag{9}$$

$$p(C_k|\varnothing) = y_k(\varnothing) = \frac{\exp(a_k)}{\sum_f \exp(a_f)} \tag{10}$$

In this study, random forest method not only used as machine learning for classification but also used as feature selection through the features importance value [38]. Features importance value in random forest method can be an aid to improving the model by understanding the significant value which contribute to high accuracy [39]. This can help shorten the training time.

KNN design used in this research is basic KNN with weighting factor of the class which can be seen in Eq. (11) [40].

$$W[c] = \frac{1}{frequency[c]} \tag{11}$$

This study used CNN as one of machine learning method which consist of convolutional layers, activation layers, pooling layers and fully connected layers. ReLu model used as activation layer on CNN. The last of fully connected layer is compute using SoftMax to define the probability of each category. CNN layers represented in

Eq. (12) to Eq. (16) [41].

$$\text{Convolutional} = z^l = h^{l-l} * W^l \tag{12}$$

$$\text{Pooling layer} = h^1_{xy} = \max_{i=0,,s,\, j=0,,s} h^{l-l}(x+i)(y+j) \tag{13}$$

$$\text{Activation layer} = ReLu(z_l) = \max(0, z_i) \tag{14}$$

$$\text{Fully connected layer} = z_l = W_l * h_{l-l} \tag{15}$$

$$\text{SoftMax} = SoftMax(z_i) = e^{zi} / \sum_j e^{zj} \tag{16}$$

where z^l is first layer pre-activation output, h^l is first activation layer, * is a discrete convolutional operator and W is the learnable parameters [41].

The feature examination conducted as follows:

1. Classified all features calculation data using the four machine learning methods to get the accuracy, precision, recall and f1 score of each method.
2. All features examined by the importance value using random forest method. Features with low importance value were removed.
3. Repeat step 1 and 2 until sufficient voice features resulted in better accuracy for voice classification.

3 Result and Discussion

Voice data acquisition result shows in Fig. 2 where Fig. 2a illustrated the PWP voice signal and Fig. 2b illustrated the healthy participants voice signal. Figure 2 shows that voice signal recorded from PWP patients have unstable phonation than healthy participants. This result is consistent with previous study [9, 20] which shows the same result. PWP shows the sign of phonation stability which can be seen in Fig. 2a as amplitude modulation [20].

Figure 3 shows the result of feature selection using random forest method. Figure 3a shows result of first run classification using random forest where features with importance value below 0.03 being removed for second run of classification. Features being removed can be seen in gray column on the Fig. 3. First random forest run removed 5 features: PPQ5 with importance value 0.039312, shimmer with importance value 0.029133, HNR with importance value 0.020867, RPDE with importance value 0.027399, Spread1 with importance value 0.027399 and correlation dimension with importance value 0.

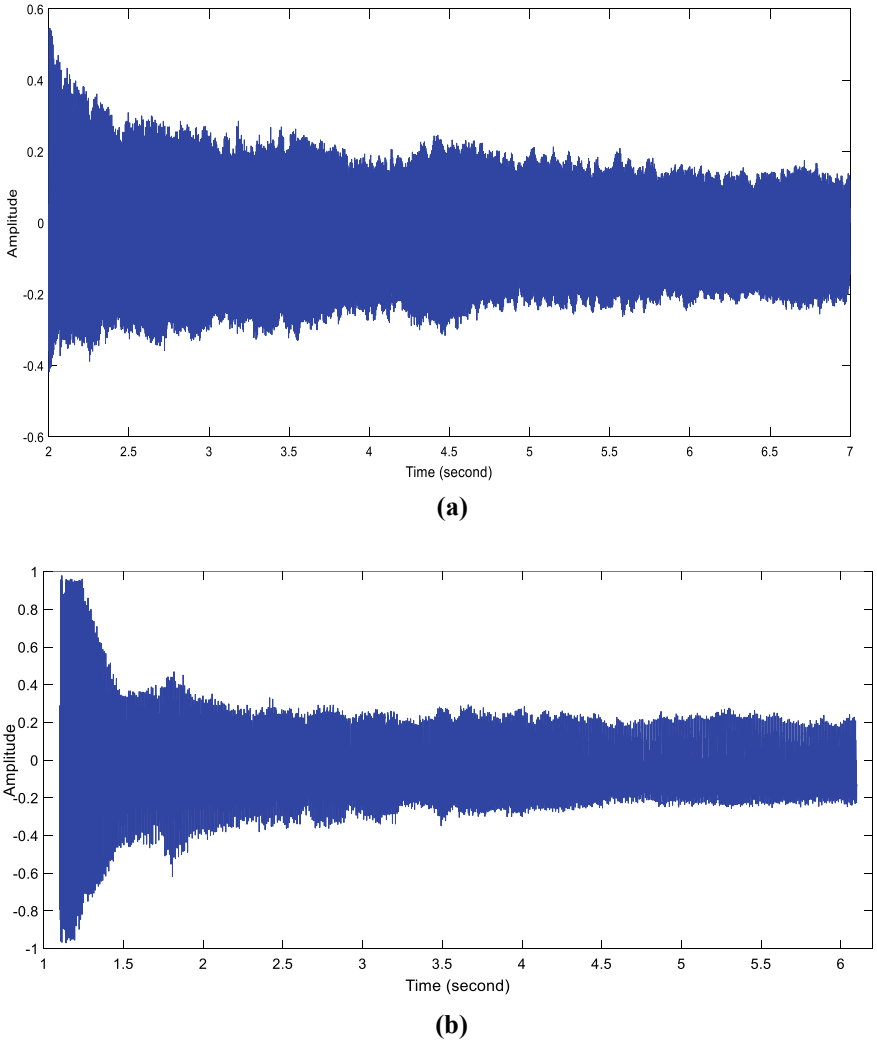
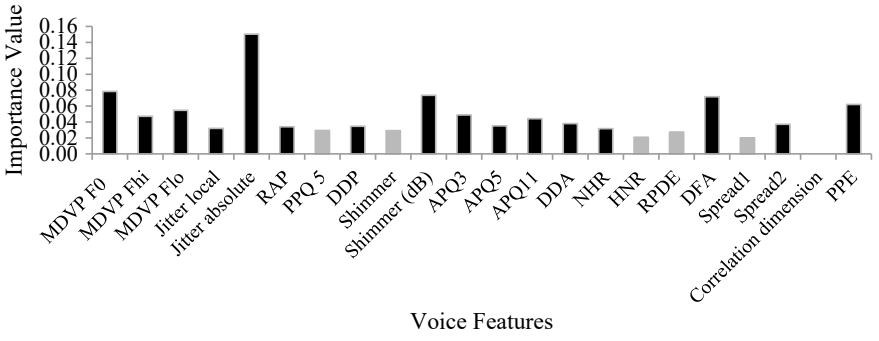


Fig. 2 Voice data acquisition result: **a** PWP participants and **b** Healthy participants

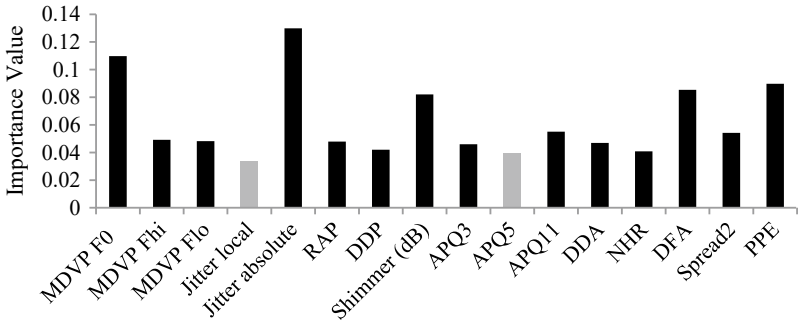
Figure 3b shows the features importance value for second run of random forest where two features with importance value below 0.04 being removed, i.e. jitter local and APQ5 each with importance value 0.033503 and 0.039727.

The third run which result represented in Fig. 3c, shows that another 3 features being removed. Those removed features were all the features with importance value below 0.05. DDP, APQ3 and DDA with importance value 0.048184, 0.047262 and 0.047704 were being removed.

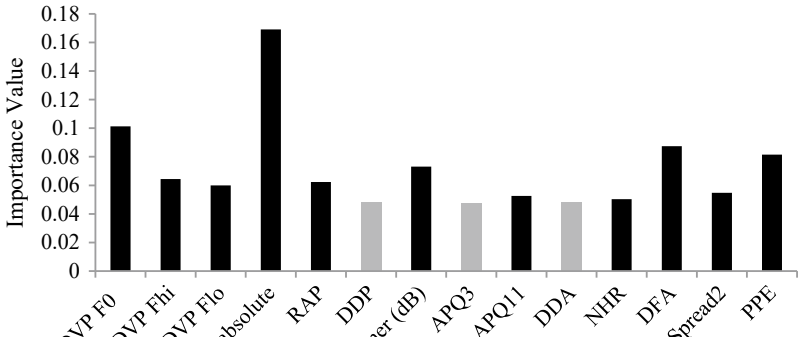
Classification result of 22 voice features using four machine learning methods, i.e. logistic regression, random forest, KNN and CNN can be seen in Table 4. Table 4



(a)



(b)



(c)

Fig. 3 Voice feature extraction result using random forest method: **a** Second run, **b** Third run and **c** Fourth run

shows the value of precision, recall, f1-score and accuracy value for each machine learning methods and class.

PD detection and monitoring studies mostly focused on binary classification, i.e. PWP-healthy. Not many researches focused on multiclassification for PD severity level detection. Some studies which focused on multiclass PD detection [23, 42–44, 45] classifies PWP into multiclass based on UPDRS rating meanwhile this present study classified PWP into four class based on Hughes scale which mostly used in Indonesia. Compared to another study which used vowel task and phonation to obtained data the logistic regression in this study and random forest method in previous study had the similar accuracy, this present study obtained 58% accuracy meanwhile another previous study obtained $52 \pm 3\%$ accuracy [23].

Table 5 shows second run result with 16 features which have been selected based on random forest method. Logistic regression, random forest and CNN shows increased accuracy result, meanwhile KNN method shows significant decreased accuracy result.

Third run classification result is shown in Table 6. Random forest method shows significant increased in accuracy meanwhile KNN shows increased accuracy compared to second run of classification. Both random forest and CNN method result in highest accuracy compared to others.

Table 7 denoted the fourth run classification result with 11 features based on Fig. 3c. Result shows that random forest and CNN method had the highest accuracy.

Table 4 Classification result using logistic regression, random forest, KNN and CNN for 22 voice features (first run)

Classifier	Class	Precision	Recall	f1-score	Accuracy (%)
Logistic regression	Possible	1.00	0.20	0.33	58
	Probable	0.70	0.54	0.61	
	Definite	0.50	1.00	0.67	
	Healthy	0.5	0.67	0.57	
Random forest	Possible	1.00	0.25	0.40	76
	Probable	0.80	0.89	0.84	
	Definite	0.50	1.00	0.67	
	Healthy	0.83	0.83	0.83	
KNN	Possible	0.75	0.75	0.75	76
	Probable	0.78	0.78	0.78	
	Definite	0.50	0.50	0.50	
	Healthy	0.83	0.83	0.83	
CNN	Possible	0.75	0.75	0.75	76
	Probable	0.78	0.78	0.78	
	Definite	0.50	0.50	0.50	
	Healthy	0.83	0.83	0.83	

Table 5 Classification result using logistic regression, random forest, KNN and CNN for 16 voice features (second run)

Classifier	Class	Precision	Recall	f1-score	Accuracy (%)
Logistic regression	Possible	0.00	0.00	0.00	52
	Probable	0.75	0.33	0.46	
	Definite	0.67	1.00	0.80	
	Healthy	0.43	1.00	0.60	
Random forest	Possible	1.00	0.25	0.40	86
	Probable	0.82	1.00	0.90	
	Definite	0.67	1.00	0.80	
	Healthy	1.00	1.00	1.00	
KNN	Possible	0.25	0.20	0.22	48
	Probable	0.75	0.46	0.57	
	Definite	0.40	1.00	0.57	
	Healthy	0.44	0.44	0.44	
CNN	Possible	1.00	0.75	0.86	95
	Probable	1.00	1.00	1.00	
	Definite	1.00	0.75	0.86	
	Healthy	0.86	1.00	0.92	

Table 6 Classification result using logistic regression, random forest, KNN and CNN for 14 voice features (third run)

Classifier	Class	Precision	Recall	f1-score	Accuracy (%)
Logistic regression	Possible	0.50	1.00	0.67	52
	Probable	0.50	0.71	0.59	
	Definite	1.00	0.33	0.50	
	Healthy	0.40	0.33	0.36	
Random forest	Possible	1.00	0.75	0.86	95
	Probable	0.90	1.00	0.95	
	Definite	1.00	1.00	1.00	
	Healthy	1.00	1.00	1.00	
KNN	Possible	0.67	0.50	0.57	57
	Probable	0.80	0.44	0.57	
	Definite	0.50	1.00	0.67	
	Healthy	0.44	0.67	0.53	
CNN	Possible	1.00	1.00	1.00	95
	Probable	1.00	1.00	1.00	
	Definite	1.00	0.75	0.86	
	Healthy	0.86	1.00	0.92	

Fewer features will contribute to lesser training time and equation code when this system implement into telehealth application.

Figure 4 shows the accuracy comparison between machine learning method used in this study. Random forest and CNN method shows that the accuracy increased along with feature selection.

Table 7 Classification result using logistic regression, random forest, KNN and CNN for 11 voice features (fourth run)

Classifier	Class	Precision	Recall	f1-score	Accuracy (%)
Logistic regression	Possible	0.50	1.00	0.67	52
	Probable	0.50	0.71	0.59	
	Definite	1.00	0.33	0.50	
	Healthy	0.40	0.33	0.36	
Random forest	Possible	1.00	0.75	0.86	95
	Probable	1.00	1.00	1.00	
	Definite	0.67	1.00	0.80	
	Healthy	1.00	1.00	1.00	
KNN	Possible	0.67	0.50	0.57	57
	Probable	0.80	0.44	0.57	
	Definite	0.50	1.00	0.67	
	Healthy	0.44	0.67	0.53	
CNN	Possible	1.00	1.00	1.00	95
	Probable	1.00	1.00	1.00	
	Definite	1.00	0.75	0.86	
	Healthy	0.86	1.00	0.92	

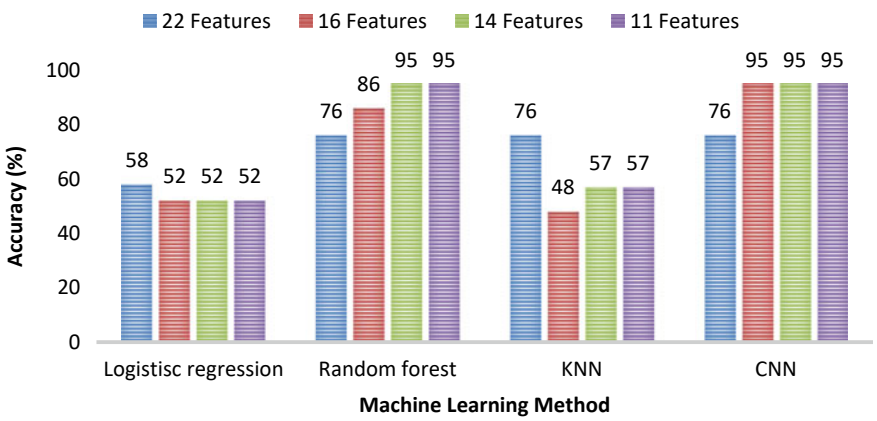


Fig. 4 Accuracy comparison between machine learning method

4 Conclusion

Voice features examination which have been conducted to understand better features with the highest accuracy for PD detection. Result shows that the most suitable result achieved with 11 features which showed in Table 7. The highest accuracy obtained with random forest and deep learning method (CNN) with 95% accuracy. This result shows similarity with previous study [38] which indicated that random forest method accuracy result tend to be increased when it is also used as feature selection technique. However the result in this study shows different accuracy for logistic regression method. Previous study shows increased accuracy tendency meanwhile this study shows decreased accuracy.

This study can be improved with another advanced voice features and compared to many other machine learning method such as support vector machine (SVM). Another data obtained task beside phonation can also be considered to improve which data collection standard procedure for PD detection an monitoring. The result of this present study can be used as telehealth application development in the future.

References

1. National Institute of Neurological Disorders and Stroke (2020) Parkinson's disease: hope through research. NIH Publication, Maryland, United States of America
2. Standaert DG, Saint-Hilaire MH, Thomas CA et al (2018) Parkinson's disease handbook. American Parkinson's Disease Association, United states of America
3. Simonet C, Schrag A, Lees AJ, Noyce AJ (2021) The motor prodromes of Parkinson's disease: from bedside observation to large-scale application. J Neurol 268(6):2099–2108
4. Bjornestad A, Tysnes OB, Larsen JP, Alves G (2016) Reliability of three disability scales for detection of independence loss in Parkinson's disease. Parkinson's Dis 2016:1–6
5. Bengacemi H, Hacine-Gharbi A, Ravier P, Abed-Meraim K, Buttelli O (2021) Surface EMG signal classification for Parkinson's disease using WCC descriptor and ANN classifier. In: De Marsico M (ed) The 10th international conference on pattern recognition applications and mathods (ICPRAM 2021). Science and Technology Publications, pp 287–294
6. Pasmanasari ED, Pawitan JA (2021) The potential of electromyography signals as markers to detect and monitor Parkinson's disease. Biomed Pharmacol J 14(1):373–378
7. Ghaheri P, Shateri A, Nasiri H (2023) PD-ADSV: an automated diagnosing system using voice signals and hard voting ensemble method for Parkinson's disease. Softw Impacts 16:1–8
8. Magesh PR, Myloth RD, Tom RJ (2020) An explainable machine learning model for early detection of Parkinson's disease using LIME on DaTSCAN imagery. Comput Biol Med 126:104041
9. Putri FT, Ariyanto M, Caesarendra W, Ismail I, Pambudi KA, Pasmanasari ED (2018) Low cost Parkinson's disease early detection and classification based on voice and electromyography signal. In: Pedrycz W, Chen S-M (eds) Computational intelligence for pattern recognition. Springer International Publishing, Cham
10. Luna-Ortiz I, Aldape-Pérez M, Uriarte-Arcia AV, Rodríguez-Molina A, Alarcón-Paredes A, Ventura-Molina E (2023) Parkinson's disease detection from voice recordings using associative memories. Healthcare 11(11):1–19
11. Alshammri R, Alharbi G, Alharbi E, Almubark I (2023) Machine learning approaches to identify Parkinson's disease using voice signal features. Front Artif Intell 6:1–8

12. Mahmood A, Mehroz Khan M, Imran M, Alhajlah O, HDhahri H, Karamat T (2023) End-to-end deep learning method for detection of invasive Parkinson's disease. Diagnostics 13(6):1–15
13. Khaskhoussy R, Ben Ayed Y (2023) Improving Parkinson's disease recognition through voice analysis using deep learning. Pattern Recognit Lett 168:64–70
14. Pah ND, Motin MA, Kumar DK (2022) Phonemes based detection of Parkinson's disease for telehealth applications. Sci Rep 12:1–9
15. Laganas C, Lakovakis D, Hadjidimitrou S, Charisis V, Dias BS et al (2022) Parkinson's disease detection based on running speech data from phone calls. IEEE Trans Biomed Eng 69(5):1573–1584
16. Bárcenas R, Fuentes-García R, Naranjo L (2022) Mixed kernel SVR addressing Parkinson's progression from voice features. Plos One 17(10):1–23
17. Dao SVT, Yu Z, Tran LV, Phan PNK, Huynh TTM, Le TM (2022) An analysis of vocal features for Parkinson's disease classification using evolutionary algorithms. Diagnostics 12(8):1–17
18. Tsanas A, Arora S (2021) Assessing Parkinson's disease speech signal generalization of clustering results across three countries: Findings in the Parkinson's voice initiative study. In: Bracken B (ed) 14th international conference on bio-inspired system and signal processing BIOSIGNAL 2021, vol 4. Science and Technology Publication, pp 124–134
19. Narendra NP, Schuller B, Alku P (2021) The detection of Parkinson's disease from speech using voice source information. IEEE/ACM Trans Audio Speech Lang Process 29(1925–1936):1
20. Majda-Zdancewicz E, Potulska-Chromik A, Jakubowski J, Nojszewska M, Kostera-Pruszczyk A (2021) Deep learning vs feature engineering in the assessment of voice signals for diagnosis in Parkinson's disease. Bull Polish Acad Sci Tech Sci 69(3):1–10. https://doi.org/10.24425/bpasts.2021.137347
21. Amato F, Borzì L, Olmo G, Orozco-Arroyave JR (2021) An algorithm for Parkinson's disease speech classification based on isolated words analysis. Health Inf Sci Syst 9(1):1–15
22. Carrón J, Campos-Roca Y, Madruga M, Pérez CJ (2021) A mobile-assisted voice condition analysis system for Parkinson's disease: assessment of usability conditions. Biomed Eng Online 20(1):1–24
23. Kodali M, Kadiri SR, Alku P (2022) Automatic classification of the severity level of Parkinson's disease: a comparison of speaking tasks, features, and classifiers. Comput Speech Lang 83(101548):1–15
24. An G, Brizan DG, Ma M, Morales M, Syed AR, Rosenberg A (2015) Automatic recognition of unified Parkinson's disease rating from speech with acoustic, i-vector and phonotactic features. In: Annual conference international speech communication and association, INTERSPEECH 2015, vol 2015, pp 508–512
25. Suppa A, Costantini G, Asci F, Di Leo P, Al-Wardat MS et al (2022) Voice in Parkinson's disease: a machine learning study. Front Neurol 13(831428):1–12
26. Little MA, McSharry PE, Hunter EJ, Spielmen J, Ramig LO (2008) Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE Trans Biomed Eng 23(1):1–7
27. Sakar CO, Serbes G, Gunduz A, Tunc HC, Nizam H et al (2019) A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform. Appl Soft Comput 74:255–263
28. Tsanas A, Little MA, McSharry PE, Ramig LO (2010) Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests. IEEE Trans Biomed Eng 57(4):884–893
29. Orozco-Arroyave JR, Arias-Londoño JD, Vargas-Bonilla JF, González-Rátiva MC, Nöth E (2014) New Spanish speech corpus database for the analysis of people suffering from Parkinson's disease. In: 9th international conference of language resources and evaluation, LREC 2014, pp 342–347
30. Rasheed J, Hameed AA, Ajlouni N, Jamil A, Özyavaş A, Orman Z (2020) Application of adaptive back-propagation neural networks for Parkinson's disease prediction. In: International conference on data analytics for business and industry: way towards a sustainable economy, bICDABI 2020, pp 1–5
31. Kadam VJ, Jadhav SM (2019) Feature ensemble learning based on sparse autoencoders for diagnosis of Parkinson's disease. Comput Commun Signal Process, 567–581

32. Martinez-Martin P, Rodriguez-Blazquez C, Forjaz MJ, Chaudhuri KR (2014) Guide to assessment scales in Parkinson's disease. Springer Healthcare, Tarporley England
33. Ma A, Lau KK, Thyagarajan D (2020) Voice changes in Parkinson's disease: what are they telling us? J Clin Neurosci 72:1–7
34. Paronen S, Vuomajoki M (2019) Voice-related changes in people with Parkinson's disease. Tampere University, Finlandia
35. Lu BY (2016) Unidirectional microphone based wireless recorder for the respiration sound. J Bioeng Biomed Sci 6(3):1–8
36. Teixeira JP, Gonçalves A (2014) Accuracy of jitter and shimmer measurements. Procedia Technol 16:1190–1199
37. Nugroho WH, Handoyo S, Akri YJ, Sulistyono AD (2022) Building multiclass classification model of logistic regression and decision tree using the chi-square test for variable selection method. J Hum Univ Nat Scieince 49(4):172–181
38. Bharath E, Rajagopalana T (2023) Parkinson's disease classification using random forest kerb feature selection. Intell Autom Soft Comput 36(2):1417–1433
39. Lewinson E (2023) Homepage. http://www.towardsdatascience.com. Accessed 17 Aug 2023
40. Xing W, Bei Y (2020) Medical health big data classification based on KNN classification algorithm. IEEE Access 8:28808–28819
41. Saputra DCE, Maulana Y, Win TA, Phann R, Caesarendra W (2023) Implementation of machine learning and deep learning models based on structural MRI for identification autism spectrum disorder. Jurnal Ilmiah Teknik Elketro Komputer dan Informatika 9(2):307–318
42. Vásquez-Correa JC, Orozco-Arroyave JR, Bocklet T, Nöth E (2018) Towards an automatic evaluation of the dysarthria level of patients with Parkinson's disease. J Commun Disord 76:21–36
43. Arias-Londoño JD, Gómez-García JA (2020) Predicting UPDRS scores in Parkinson's disease using voice signals: a deep learning/transfer-learning-based approach. In: Automatic assessment of Parkinsonian speech, pp 100–123
44. Arias-Vergara T, Vásquez-Correa JC, Orozco-Arroyave JR (2017) Parkinson's disease and aging: analysis of their effect in phonation and articulation of speech. Cogn Comput 9(6):731–748
45. Rahman A, Khan A, Raza AA (2020) Parkinson's disease detection based on signal processing algorithms and machine learning. CRPASE: Trans Electron Comput Eng 06(03):141–145

Bone Drilling Vibration Signal Classification Using Convolutional Neural Network to Determine Bone Layers

Wahyu Caesarendra, Putri Wulandari, Kamil Gatnar, and Triwiyanto

Abstract In orthopedic surgery, the bone drilling task is the main factor for the successful completion of the surgery. The bone drilling task depends on a high level of dexterity and experience of the orthopedist and surgeon. It is because the bone drilling resistance is relatively large and sometimes violent vibrations might cause difficulties in grasping the hand-piece. In a worse case, it might even break the slender drill. The objective of this paper is to gain an understanding of the frequency properties of bone in order to improve visualization and training. These properties can be detected using different imaging methods or techniques for processing signals. The experimental setup includes the robotic arm to provide an accurate thickness layer and consistent penetration of the drilling. Three-axis accelerometer equipped with National Instrument data acquisition (DAQ) was used in the experiment to acquire the vibration signal on different bone layers. This study proposes a successful approach to categorize bone drilling levels using a Convolutional Neural Network (CNN) with a customized architecture designed for this purpose. The CNN is utilized to classify the raw vibration signal into three distinct labels or layers, namely periosteum, first cortical, and spongy. The results of the study indicate that the CNN can accurately classify the three bone layers, with a higher degree of accuracy for the periosteum and first cortical layers, achieving over 98% accuracy, and a 100% accuracy for the spongy layer. This is due to the unique vibration signal of the spongy layer, which differs from the other two layers.

Keywords Bone drilling · Bone layer classification · CNN · Vibration signal

W. Caesarendra (✉) · P. Wulandari
Faculty of Integrated Technologies, Universiti Brunei Darussalam, Jalan Tungku Link, Gadong 1410, Brunei
e-mail: wahyu.caesarendra@ubd.edu.bn

W. Caesarendra · K. Gatnar
Faculty of Mechanical Engineering, Opole University of Technology, 76 Proszkowska St., 45-758 Opole, Poland

Triwiyanto
Department of Electro Medic Engineering, Health Polytechnic Ministry of Health Surabaya, Surabaya, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_40

1 Introduction

Bone drilling is a frequently performed procedure in orthopedic surgery, including total knee replacement surgery, fracture treatment, and dental treatments [1]. During drilling surgery, it is crucial to drill both the outer cortical bone and the softer cancellous bone inside. However, if the drilling parameters are not optimized, the brittle nature of the bone can lead to the generation of significant axial force and heat during the drilling process [1]. In bone drilling, the three layers of the inner base bone—first cortical bone, spongy bone, and second cortical bone layer—are typically used to analyze and treat bone structures with nails and screws [2, 3].

Previous studies have shown that micro-damages in surrounding bone and damage to nerves are in the self-drilling technique greater than self-tapping technique (drilling with a pilot hole) and histomorphometric parameters like length, number, and density of cracks are much greater and undesirable in the self-drilling technique [4, 5]. As the drilling process takes place, there is an increase in temperature, which can lead to bone necrosis [6]. Bone necrosis refers to the irreversible death of bone cells when the temperature surpasses a certain threshold. Hence, employing the drilling machine for implementing the self-tapping technique during orthopedic procedures is crucial and necessary.

Currently, in orthopedic surgery, bone drilling is carried out using manual hand drills, and the speed at which the drill bit moves is controlled by the surgeon. The effectiveness of the drilling largely relies on the surgeon's skill and their ability to gauge the drilling process based on their perception [7, 8]. The force experienced by the surgeon during drilling is subjective, as it depends on factors such as the speed of the drill bit, the condition of the bone, and the type of drill bit employed [9, 10].

To tackle this issue, a new drilling scheme—sensor-aided vibrational bone drilling was introduced with the purpose of reducing cutting forces and heat generation. In this investigation, accelerometer sensors are applied to a drill bit to acquire a vibration signal. In [11] an experimental study on bovine bone utilizing ultrasonically assisted drilling discovered that force and torque were significantly reduced as compared to the conventional drilling approach. In addition, Gupta et al. [7, 12] discovered that compared to conventional drills, sensor-aided drilling could create lower temperatures when the vibration frequency was held at 20 kHz and the vibration amplitude was kept at 4–20 μm.

In a study conducted by [13], a similar approach using a CNN was used to classify bone drilling vibration signals into two different layers: cortical and cancellous. This study uses a combination of time-domain and frequency-domain features to train a CNN and achieves high classification accuracy for both layers. Overall, current studies demonstrate the potential of CNN-based methods for accurately classifying bone drilling vibration signals and determining bone layers. For this reason, this study uses CNN as the main method.

This paper aims to know the frequency characteristics of bone for visualization and training. The frequency characteristics of bone layer classification refer to the distinctive patterns or properties exhibited by different layers of bone when analyzed

in terms of frequency. These characteristics can be observed through various imaging techniques or signal processing methods.

By analyzing the frequency response of bone layers, it is possible to differentiate between cortical bone and cancellous bone based on their unique frequency patterns [14]. Cortical bone typically exhibits higher frequency responses, indicating higher density and hardness, while cancellous bone displays lower frequency responses due to its porous nature. Understanding the frequency characteristics of bone layers can aid in various applications, including bone quality assessment, orthopedic surgeries, and the development of diagnostic tools for bone-related conditions.

2 Materials and Methods

2.1 Convolutional Neural Network (CNN)

Deep learning has become a hot topic in the world of machine learning owing to its significant capabilities in modeling various complex data, such as images and sounds. The Deep Learning method that currently has the most significant results in image recognition is the Convolutional Neural Network (CNN) [14, 15]. A convolutional neural network (CNN) is a deep learning method that has been widely applied to image data. CNN is a type of feedforward ANN algorithm commonly used to process data based on a grid structure [16]. To process the data, the CNN used several arrays as inputs. To achieve satisfactory results, a CNN must be designed with a more complex architecture. CNN is usually used for image data; however, previous research presents the application of CNN for 1-D data or vibration data alone [17]. To utilize CNN for image data, larger-pixel images are required to achieve an improved quality of acquisition. Consequently, training a CNN requires significant computational resources and faces implementation constraints owing to its slow processing speed. At this time, several algorithms to increase CNN speed and optimization have been explored more deeply and have begun to be widely researched [18, 19].

There are several models and architectures in the training process of CNN including ReLu (rectified linear unit), Dropout, and batch normalization, VGG-16 (16 layers), LeNet, googleNet (22 layers), AlexNet (8 layers), ResNet (152 layers), etc. [20]. These models and architectures have merits and demerits and can be implemented in different problems. This is because the model and architecture are used in accordance with the state of the object of identification in different layers of CNN of the training stage [21, 22]. The modern CNN namely LeNet-5 was presented by Tavakkoli et al. [23]. LeNet-5 has 7 layers excluding the input layer including C1, S2, C3, S4, C5, and F6 output. In addition, Zhang et al. [21], reported that not only a number of layers in the CNN model which is necessary to be considered, but also its configuration, feature, and parameter size.

Each training model has a unique layer according to the type of architecture utilized during training. The layers of the CNN are as follows:

- Sub-sampling layer: This layer is frequently employed to scale back the input to its original dimensions. The CNN sub-sampling layer aims to minimize the number of parameters required. The success of CNN could be accelerated by this.
- Convolutional layer: works by evaluating filters from input images and mimicking the characteristics of the visual cortex in the brain. The filters learned in this layer can take various forms; for example, if they are used to learn from an image, these filters might learn edge detection. The following is the equation for the convolutional layer [24]:

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau \tag{1}$$

- Loss layer: The CNN output, which is a crucial component of the neural network, is the loss layer. To update weights, this function computes gradients.
- Fully connected layer: The majority of neural networks contain a fully linked layer. The output from fully connected layers is obtained via matrix multiplication.
- ReLu layer: The activation function in neural networks is similar to or utilized for thresholding in the ReLu layer. The equation for the ReLu layer max(0, x) using the sigmoid function is as follows.
- Pooling layer: Similar to the convolutional layer, this layer is important for reducing the spatial size of convolution features. To reduce computing power, a process for reducing the dimensions is required. This layer is useful for extracting rotating and invariant domain features so that the model training process becomes more effective. The pooling layer was divided into two parts: max pooling and min pooling.
- Flattening layer: The flattening layer transforms the shape of the matrix data into a one-dimensional array that is fed into the next layer. This layer is connected to the final classification model, known as the fully connected layer. It arranges all the pixel data into a single row and establishes connections with the last layer.

The previously mentioned algorithm is well-suited for addressing challenges in tasks such as image classification, detecting objects, determining object positions, and segmenting images [25]. Examples of case studies on image recognition using CNN usually have three stages: input, CNN, and output. A detailed description of the CNN stages is presented in [23, 26, 27].

2.2 CNN Structure of the Present Study

The process of classifying bone drilling levels, as depicted in Fig. 1, is a three-stage procedure that includes input, convolution, and classification. The sensor utilized in this study is an accelerometer that provides output along three axes: the x-axis, y-axis, and z-axis. The Convolutional Neural Network (CNN), the classifier used in

this research, has an architecture as shown in Fig. 1. It comprises six layers: two one-dimensional convolution layers, a max pooling layer, a global average pooling layer, a dropout layer, and a fully connected layer which also serves as the output layer. The CNN accepts raw data signals from the accelerometer across three axes. The convolution process involves kernels of dimension 4×1 interacting with the three raw accelerometer signals, each of dimension 100×1. This convolution process on the first layer yields a one-dimensional output of 97×1. Following this, a max pooling process is applied which reduces the dimension to 24×1. This output then undergoes another convolution process with a 4×1 kernel. The global average pooling layer is implemented to further reduce the node dimensions. This is crucial for managing the computational complexity of the model. To prevent overfitting and increase the robustness of the system, the output from the Global Average Pooling layer is passed through a dropout layer. Overfitting occurs when a model learns the training data too well, to the point where it performs poorly on unseen data. The dropout layer mitigates this by randomly setting a fraction of input units to zero at each update during training time.

The final layer in this architecture is the fully connected layer. It consists of a set of nodes with three output nodes corresponding to the three classes. This layer performs classification on the features extracted by the preceding layers. The complete parameter details of this neural network model are provided in Table 1. These parameters include information such as the number and type of layers in the model, their configuration, and how they interact with each other. In conclusion, this research presents an effective method for classifying bone drilling levels using a Convolutional Neural Network with a specific architecture tailored for this task. The use of an accelerometer sensor for data collection and various layers for feature extraction and classification demonstrates an innovative approach to leveraging machine learning techniques in medical research.

2.3 Experimental Setup

This study uses a 3-axis accelerometer with National Instrument (NI) data acquisition (DAQ) in the bone drilling experiment. The experiment was conducted using a Dobot Magician Robot with an attached hand drill on an end effector part of the Dobot, as presented in Fig. 2. The hand drill was rotated at 500 rpm. Bone drilling was performed intermittently from layer 1 (periosteum), and layer 2 (first cortical) until the drill bit was in layer 3 (spongy). In each layer, the vibration signal was acquired with a feed rate of about 0.002–0.006 in/min. The three different depths represent three different layer-level classes, as described in Table 2 [27]. Table 2 also presents the illustration of bone layers used in the experiment associated with the thickness for each layer.

Two triaxial accelerometers type 4535-B from Brüel and Kjær were used in the experiment. One accelerometer was attached near the hand drilling and another one was attached to the bone that will undergo drilling to obtain vibration signals during

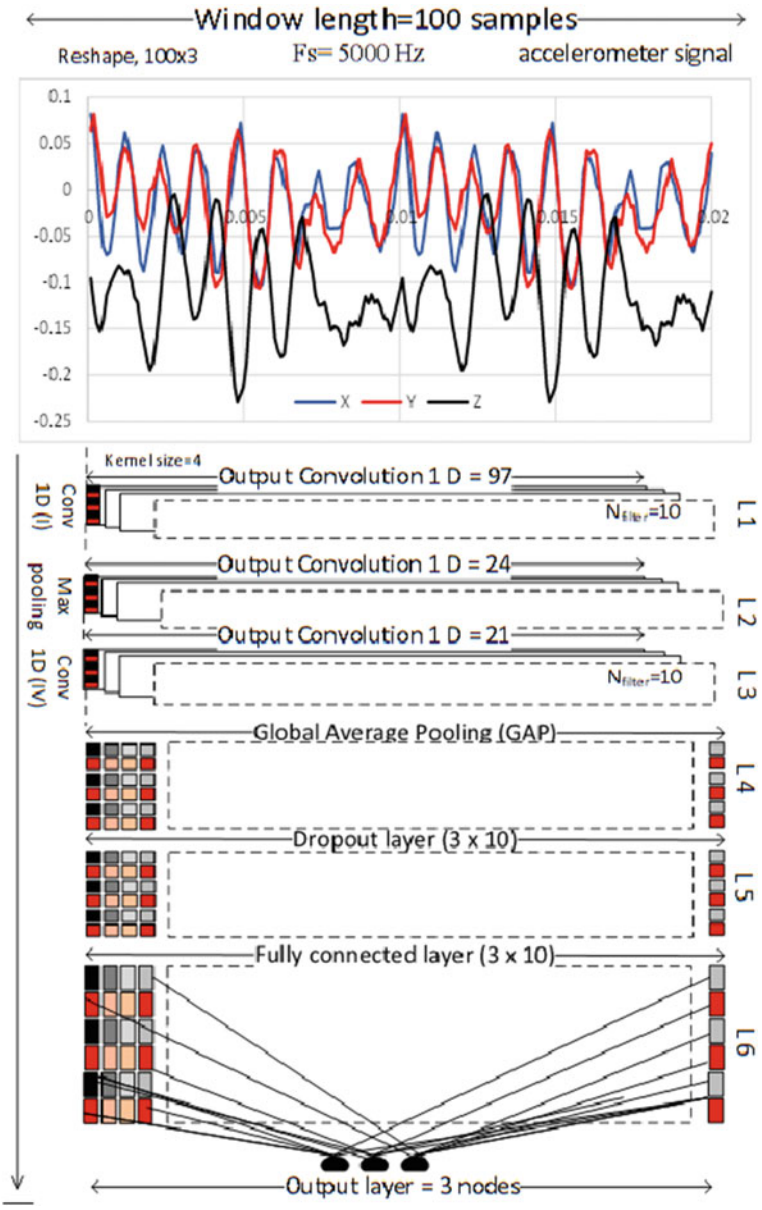


Fig. 1 CNN model and architecture for vibration signals from 3 bone layers

Table 1 Neural network model

Layer (type)	Output shape	Parameter
Reshape_8 (Reshape)	(None, 100, 3)	0
conv1d_14 (Conv1D)	(None, 97, 10)	130
max_pooling1d_7 (MaxPooling 1D)	(None, 24, 10)	0
conv1d_15 (Conv1D)	None, 21, 10)	410
global_average_pooling1d_8	(None, 10)	0
dense_8 (Dense)	(None, 3)	33

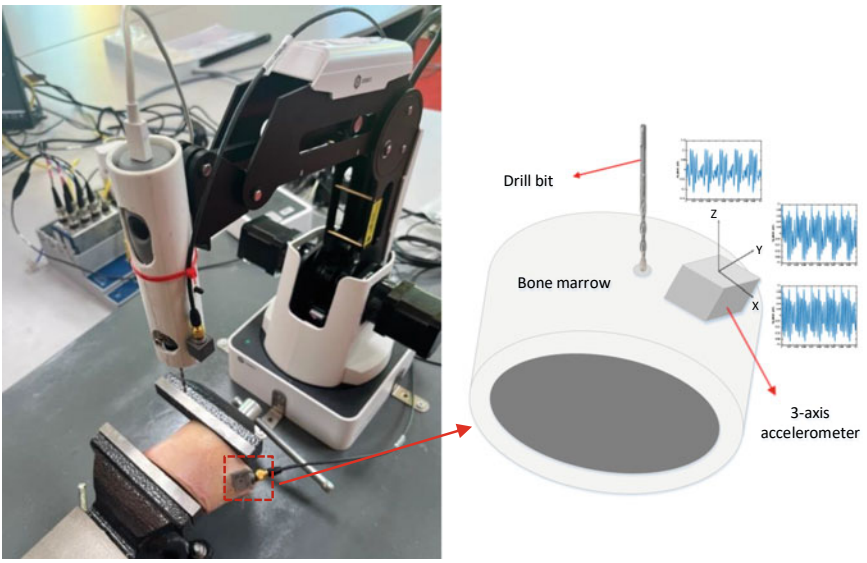
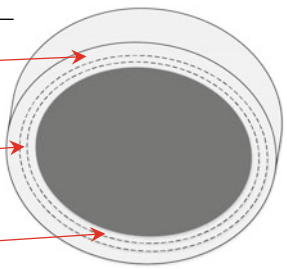


Fig. 2 Bone drilling experiment and sensor position

Table 2 Bone layer classification

Layer	Depth (mm)
Periosteum	2 – 3
First cortical	0.7 – 0.86
Spongy	1.09 – 2.12



the drilling process, as shown in Fig. 2. The NI (DAQ) data acquisition module and LabVIEW program were used to acquire vibration signals in a digital format with a sampling frequency of 5 kHz. The signals corresponding to the accelerometer sensor were synchronized offline. The Dobot Magician robot is used to assist hand drilling in performing constant drilling. The objective of using the Dobot Magician robot is to control the drilling process and to provide an accurate depth penetration for each layer in the bone drilling experiment. Note that this study only presents the vibration signal analysis from the accelerometer attached to the bone.

The experimental setting is problematic when reviewing the literature. Experiments are rarely performed intraoperatively in humans and are mostly performed in animal models or cadavers. Porcine and dog bones have been shown to be most similar to human samples [27]. In this study, fresh bovine femur of cow bone was used as an experimental material because its thermal properties are similar to those of human bones. First, soft tissue was removed to reveal the femoral diaphysis. Second, the marrow of the femoral diaphysis was removed. During drilling, the feed rate and displacement of the bone drill were controlled automatically through the Dobot Magician robot coding using Blokly. The surgical drill bit diameter was within the range of 3.5–4.5 mm. Vibration data were collected and stored in LabVIEW at a sampling frequency of 5 kHz. Data post-processing was performed using MATLAB. An example of raw vibration signals of hole #5 for layer 1, 2, and 3 on LabVIEW are presented in Fig. 3.

3 Results

3.1 Vibration Signal of Bone Layers

The vibration data of the bone drilling experiment was presented in Figs. 4, 5 and 6. Figure 4 presents one-second vibration data of the x-axis accelerometer. Figure 4a, b, and c are the vibration signals from the periosteum layer, first cortical layer, and spongy layer, respectively. It is visually seen that the raw vibration signals of different layers are difficult to distinguish. However, if the vibration signal was plotted only for 0.2 s (0.5–0.7 s), it shows a different vibration signal shape as presented in zoom in parts of Fig. 4a–c.

Figure 5 presents one-second vibration data of the y-axis accelerometer. Figure 5a, b, and c is the vibration signal from the periosteum layer, first cortical layer, and spongy layer, respectively. Similar to Fig. 4, the raw vibration signals of different layers in Fig. 5 are difficult to distinguish. However, if the vibration signal was plotted only for 0.2 s (0.5–0.7 s), it shows a different vibration signal shape as presented in zoom in parts of Fig. 5a–c.

Figure 6 presents one-second vibration data of the z-axis accelerometer. Figure 6a, b, and c are the vibration signals from the periosteum layer, first cortical layer, and spongy layer, respectively. Similar to Figs. 4 and 5, the raw vibration signals of

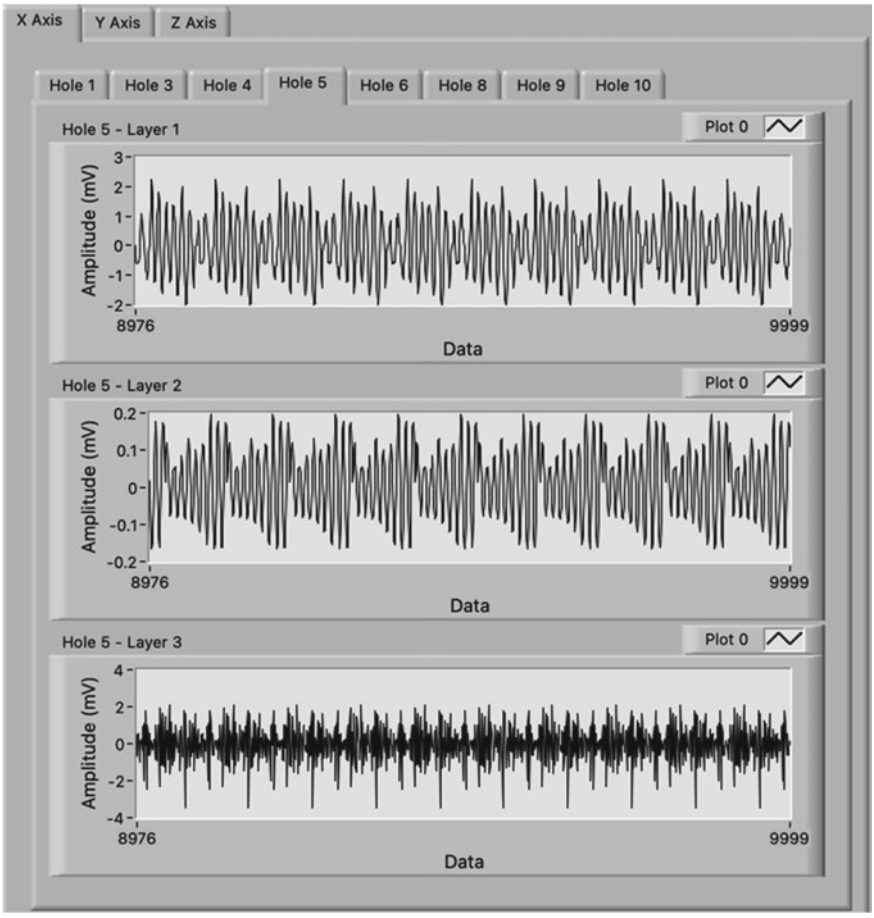


Fig. 3 Raw vibration signal on LabVIEW from x-axis, hole #5

different layers in Fig. 6 are difficult to distinguish. However, if the vibration signal was plotted only for 0.2 s (0.5–0.7 s), it shows a different vibration signal shape as presented in Zoom in parts of Fig. 6a–c.

In addition to the different vibration signals for each bone layer according to Figs. 3, 4 and 5, another obvious feature is the increasing vibration signal amplitude from the periosteum, first cortical, to the spongy layer. A detailed of the vibration amplitude is represented in the root mean square (RMS) of each vibration signal as presented in Table 3. It is shown that the RMS amplitude increases as the drilling process goes deeper. It indicates that the bone structure of each layer is different and when the bone structure has contact with the drill bit with predetermined RPM, it triggers a different vibration characteristic. In addition, an increase in RPM values from the periosteum to the first cortical layer is not as high as an increase in RPM

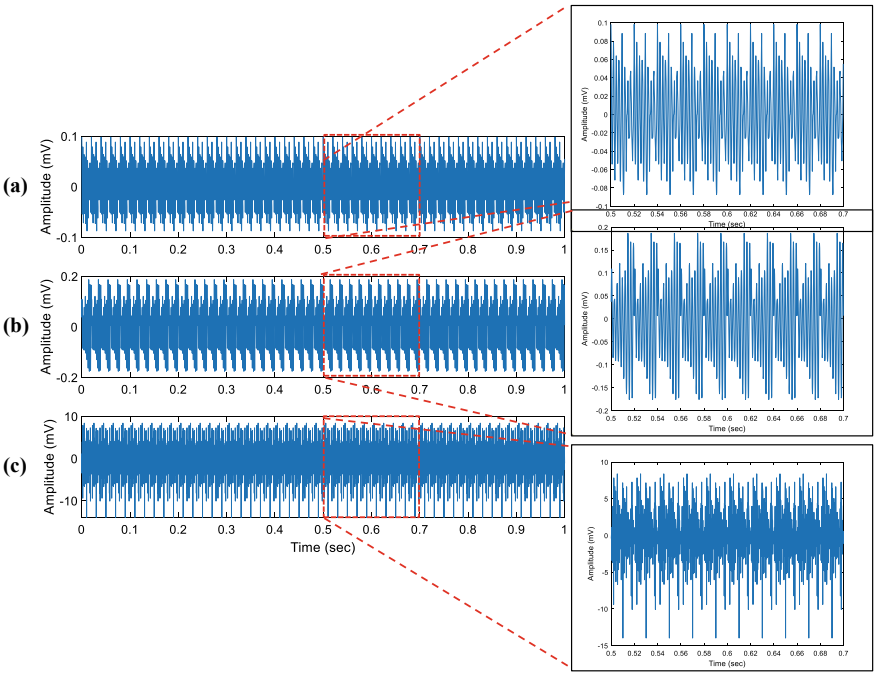


Fig. 4 Bone vibration signal from x-axis accelerometer: **a** periosteum layer; **b** first cortical layer; **c** spongy layer

values from the first cortical to the spongy layer. It indicates that the spongy layer has less stiffness than the other two layers (periosteum and first cortical).

3.2 CNN Classification

The confusion matrix shown in Fig. 6 illustrates the accuracy of the classifier for each class (periosteum, first cortical, and spongy). For the periosteum class, the classifier was able to classify 996 data points correctly, but it misclassified one data point each for the first cortical and spongy classes. In the first cortical class, the classifier correctly classified 994 data points. However, there were six instances of misclassification: the classifier incorrectly identified two data points as periosteum and four data points as spongy. For the spongy class, the classifier correctly classified all 997 data points with no instances of misclassification. These results highlight the effectiveness of the classifier in distinguishing between the three classes. The high accuracy rates suggest that the classifier has learned meaningful patterns from the data that allow it to make accurate predictions. However, the few instances of misclassification indicate areas where the classifier may be improved. For example, additional training or

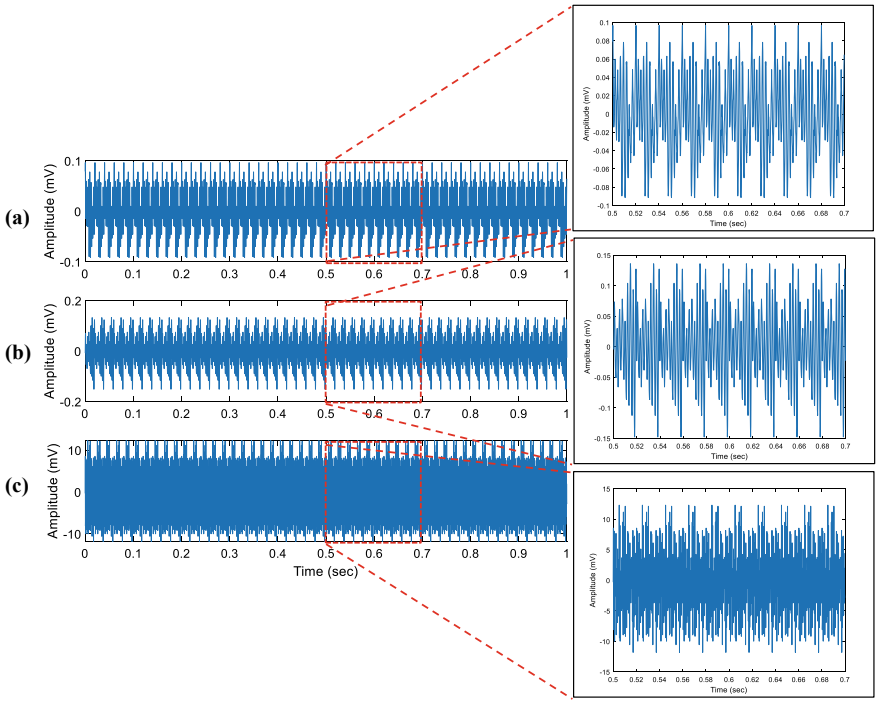


Fig. 5 Bone vibration signal from y-axis accelerometer: **a** periosteum layer; **b** first cortical layer; **c** spongy layer

fine-tuning could potentially reduce these errors. Alternatively, further investigation into these misclassified instances may reveal specific characteristics or anomalies that could inform adjustments to the model or pre-processing steps. Overall, while there is room for improvement, the classifier demonstrates strong performance in this classification task.

According to the confusion matrix presented in Fig. 7, the classification of the periosteum and first cortical layer has misclassification, whereas the spongy class does not have misclassification. In detail, two data points in the periosteum class have been predicted as first cortical (1 data point) and spongy (1 data point). This indicates that a few of the vibration signal of the periosteum layer has similarity to the first cortical and spongy layer. In addition, six data points of the first cortical have been predicted as periosteum (2 data points) and spongy (4 data points). Apart of similarity the vibration signal to other layers, the 4 misclassified data points of the first cortical layer indicate that the vibration signal is in the transition between the first cortical to the spongy layer.

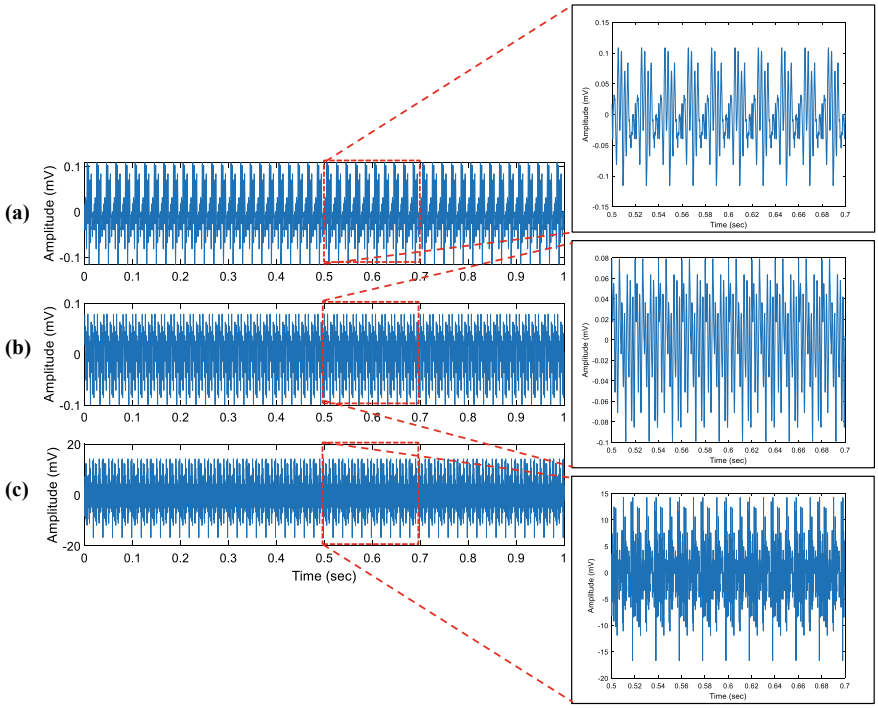


Fig. 6 Bone vibration signal from z-axis accelerometer: **a** periosteum layer; **b** first cortical layer; **c** spongy layer

Table 3 Root mean square (RMS) value of bone vibration signal

Layer	x-axis (mV)	y-axis (mV)	z-axis (mV)
Periosteum	0.044	0.044	0.044
First cortical	0.092	0.063	0.051
Spongy	4.044	5.505	5.853

4 Discussion

In addition to the confusion matrix result presented in Fig. 3, Table 4 represents the performance metrics of a classifier model. The model has been evaluated on three classes, labeled 0, 1, and 2. For class 0, the model achieved a precision, recall, and F1-score of 1. This indicates that the model perfectly identified all instances of class 0 without any false positives or false negatives. For class 1, the model also achieved a precision and F1-score of 1 but had a slightly lower recall of 0.99. This suggests that the model was excellent at predicting class 1 but missed a small fraction of actual class 1 instances. For class 2, similar to class 0, the model achieved perfect scores across all metrics. The overall accuracy of the model was 1, indicating that it

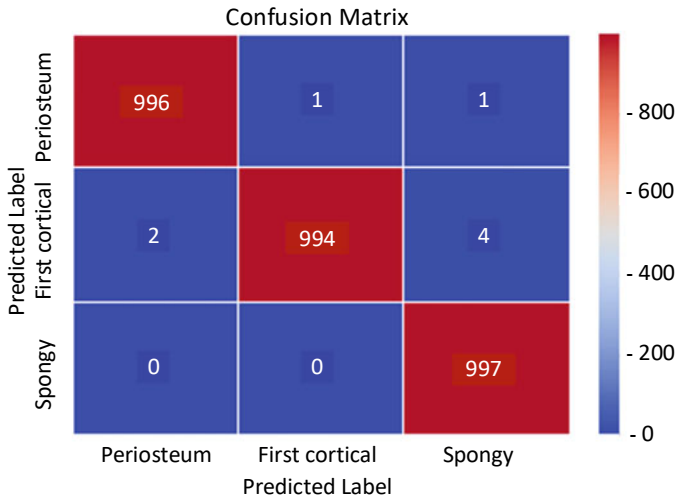


Fig. 7 Confusion matrix of bone classification

correctly classified all instances in the majority of cases. The excellent performance of the model is demonstrated by achieving perfect scores for the macro-average and weighted-average metrics calculated across all classes. These results reaffirm the model's outstanding overall capabilities and ability to excel consistently across a wide range of classification tasks.

The confusion matrices for three classes i.e. periosteum, first cortical, and spongy layers offer insightful information regarding how the classifier performs. These results highlight the classifier's ability to accurately distinguish between these classes, with high levels of accuracy being clear from the findings of the study.

However, the misclassifications of the periosteum and first cortical layer imply that the classifier occasionally has trouble identifying fine distinctions. The probable overlap or closeness of the vibration signal patterns between these layers can be explained by this misclassification, particularly when a transition from the first cortical layer to the spongy layer occurs. These errors may decrease with greater study and classifier improvement (via training or tweaking). Additionally, a more

Table 4 Classification report for test data

Class	Precision	Recall	*F*1-score	Support
0	1	1	1	998
1	1	0.99	1	1000
2	1	1	1	997
Accuracy			1	2995
Macro avg	1	1	1	2995
Weighted avg	1	1	1	2995

thorough analysis of these incorrectly classified occurrences might reveal specific traits or abnormalities that could guide model alterations or preprocessing methods.

The categorization results, which were largely accurate, showed that it was successful in extracting significant patterns from the data. These results support ongoing work to improve and expand the classifier capabilities in this particular domain.

To lessen the effect of drilling on bone vibrations and bone damage, this discovery will be highly beneficial to surgeons undergoing bone surgery. The performance of this classification model was outstanding across all the measures and classes. It reliably distinguishes each class with a minimum number of errors, making it an effective tool for this particular classification assignment.

5 Conclusions

A classification study of bone layers based on vibration signal and CNN method has been presented. The vibration signal was acquired using accelerometer and LabVIEW from the bone drilling experiment. This preliminary study used only limited vibration data (hole #5) for CNN classification. For Class 1, the model attained a precision and F1-score of 1, with a marginally lower recall of 0.99. This indicates the model's exceptional ability to predict class 1 instances, although it missed a small portion of the actual class 1 cases. Regarding class 2, akin to class 0, the model achieved flawless scores for all metrics. The overall accuracy of the model reached 1, signifying its accurate classification of the majority of instances. Both the macro-average and weighted-average scores, encompassing all classes, were also impeccable, showing the model's outstanding performance. According to the CNN results the three bone layers can be classified with a higher degree of accuracy. For the periosteum and first cortical layer, the accuracy is higher than 98%. For the spongy layer, it achieved a 100% accuracy. This is because the vibration signal of the spongy layer is different than the other two layers. In future works, more data will be acquired to test the reliability result of CNN and other classifiers will be tested.

Acknowledgements The authors also acknowledge the Polish National Agency for Academic Exchange (NAWA) No. BPN/ULM/2022/1/00139/U/00001 for financial support.

References

1. Hassanalideh HH, Gholampour S (2020) Finding the optimal drill bit material and proper drilling condition for utilization in the programming of robot-assisted drilling of bone. CIRP J Manuf Sci Technol 31:34–47. https://doi.org/10.1016/j.cirpj.2020.09.011
2. Mayer M, Lin HH, Peng YH, Lee PY, Wang ML (2014) A drill signal detection technology for handheld medical drilling device. In: 2014 international symposium on computer, consumer and control, pp 958–961. https://doi.org/10.1109/IS3C.2014.251

3. "Anatomy & Physiology," https://open.oregonstate.education/aandp/chapter/6-3-bone-structure, September 22, 2022
4. Shakouri E, Haghighi Hassanalideh H, Gholampour S (2018) Experimental investigation of temperature rise in bone drilling with cooling: a comparison between modes of without cooling, internal gas cooling, and external liquid cooling. Proc Inst Mech Eng H 232(1):45–53. https://doi.org/10.1177/0954411917742944
5. Viguet-Carrin S, Garnero P, Delmas PD (2006) The role of collagen in bone strength. Osteoporos Int 17(3):319–336. https://doi.org/10.1007/s00198-005-2035-9
6. Matthews AH, Davis DD, Fish MJ, Stitson D (2022) Avascular necrosis. StatPearls Publishing
7. Gupta V, Pandey PM, Silberschmidt VV (2017) Rotary ultrasonic bone drilling: improved pullout strength and reduced damage. Med Eng Phys 41:1–8. https://doi.org/10.1016/j.medengphy.2016.11.004
8. Brioschi V, Cook J, Arthurs GI (2016) Can a surgeon drill accurately at a specified angle? Vet Rec Open 3(1):e000172. https://doi.org/10.1136/vetreco-2016-000172
9. Pandey RK, Panda SS (2013) Drilling of bone: a comprehensive review. J Clin Orthop Trauma 4(1):15–30. https://doi.org/10.1016/j.jcot.2013.01.002
10. Han Y, Lv Q, Song Y, Zhang Q (2022) Influence of parameters on temperature rise and chips morphology in low-frequency vibration-assisted bone drilling. Med Eng Phys 103:103791. https://doi.org/10.1016/j.medengphy.2022.103791
11. Alam K, Mitrofanov AV, Silberschmidt V (2009) Measurements of surface roughness in conventional and ultrasonically assisted bone drilling. Am J Biomed Sci 2009:312–320. https://doi.org/10.5099/aj090400312
12. Gupta V, Pandey PM (2016) Experimental investigation and statistical modeling of temperature rise in rotary ultrasonic bone drilling. Med Eng Phys 38(11):1330–1338. https://doi.org/10.1016/j.medengphy.2016.08.012
13. Tran T, Pham NT, Lundgren J (2022) A deep learning approach for detecting drill bit failures from a small sound dataset. Sci Rep 12(1):9623. https://doi.org/10.1038/s41598-022-13237-7
14. Gautam D, Rao KPV (2021) Nondestructive evaluation of mechanical properties of femur bone. J Nondestr Eval 40. https://doi.org/10.1007/s10921-021-00754-0
15. Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. Insights Imaging 9(4):611–629. https://doi.org/10.1007/s13244-018-0639-9
16. Fukushima K (1980) Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 36(4):193–202. https://doi.org/10.1007/BF00344251
17. Lindsay G (2020) Convolutional neural networks as a model of the visual system: past, present, and future. J Cogn Neurosci 33:1–15. https://doi.org/10.1162/jocn_a_01544
18. Cao X-C, Chen B-Q, Yao B, He W-P (2019) Combining translation-invariant wavelet frames and convolutional neural network for intelligent tool wear state identification. Comput Ind 106:71–84. https://doi.org/10.1016/j.compind.2018.12.018
19. Habib G, Qureshi S (2022) Optimization and acceleration of convolutional neural networks: a survey. J King Saud Univ Comput Inf Sci 34(7):4244–4268. https://doi.org/10.1016/j.jksuci.2020.10.004
20. Das S (2017) CNN architectures: LeNet, AlexNet, VGG, GoogLeNet, ResNet and more. Analytics Vidhya 17
21. Zhang Q, Zhang M, Chen T, Sun Z, Ma Y, Yu B (2019) Recent advances in convolutional neural network acceleration. Neurocomputing 323:37–51. https://doi.org/10.1016/j.neucom.2018.09.038
22. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv 1409.1556
23. Shrestha A, Mahmood A (2019) Review of deep learning algorithms and architectures. IEEE Access 7:53040–53065. https://doi.org/10.1109/ACCESS.2019.2912200
24. Cui N (2018) Applying gradient descent in convolutional neural networks. J Phys Conf Ser 1004:012027. https://doi.org/10.1088/1742-6596/1004/1/012027

25. Guo Y, Liu Y, Georgiou T, Lew MS (2018) A review of semantic segmentation using deep neural networks. Int J Multimed Inf Retr 7(2):87–93. https://doi.org/10.1007/s13735-017-0141-z
26. Nguyen PT, Di Ruscio D, Pierantonio A, Di Rocco J, Iovino L (2021) Convolutional neural networks for enhanced classification mechanisms of metamodels. J Syst Softw 172:110860. https://doi.org/10.1016/j.jss.2020.110860
27. Eriksson AR, Albrektsson T, Albrektsson B (1984) Heat caused by drilling cortical bone. Temperature measured in vivo in patients and animals. Acta Orthop Scand 55(6):629–631. https://doi.org/10.3109/17453678408992410

Enhancing Accuracy in the Detection of Pneumonia in Adult Patients: An Approach by Using Convolutional Neural Networks



Lusiana, Arvita Agus Kurniasari, and Izza Fahma Kusumawati

Abstract Pneumonia is a significant respiratory condition that can cause severe health complications in adult patients if not diagnosed and treated promptly. This study focuses on enhancing the accuracy of pneumonia detection in adult patients using Convolutional Neural Networks (CNNs). The primary objective of this research is to improve upon the previous accuracy rate of 89% achieved by prior studies. A CNN architecture with five hidden layers was designed and implemented to pursue this goal. The dataset consisted of 5,000 medical images used for training, testing, and random checks to assess the model's performance comprehensively. In this research, we compared various compositions to achieve the highest precision. Changes in the number of image data with a normal data value of 1800 and pneumonia data of 3200 achieve the highest degree of precision, namely 0.9167. The accuracy of changes to the epoch value 20 is 0.9721. Using a size of 244 × 244 yields a precision of 0.9443. The proposed CNN architecture significantly advanced pneumonia detection accuracy through meticulous data preprocessing, augmentation, and training. The experimental results showed a notable improvement in accuracy, reaching a value of 91%. This enhancement can be attributed to the depth and complexity of the CNN architecture, along with the augmentation techniques employed. The achieved accuracy indicates the potential of CNNs in aiding medical professionals with more reliable and accurate pneumonia diagnoses in adult patients. The findings of this study underscore the efficacy of using CNNs in medical image analysis, specifically for pneumonia detection. Further exploration of larger datasets, more intricate architectures, and incorporation of additional diagnostic features could potentially lead to even higher accuracy rates, contributing to improved patient care and diagnostic outcomes.

Lusiana (✉)
Department of Medical Electronics Technology, Poltekkes Kemenkes Surabaya, Surabaya, Indonesia
e-mail: lusiana.tekmed@poltekkesdepkes-sby.ac.id

A. A. Kurniasari
Department of Information Technology, Politeknik Negeri Jember, Jember, Indonesia

I. F. Kusumawati
Department of Electrical Engineering, Universitas Muhammadiyah Surabaya, Surabaya, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_41

Keywords Pneumonia · Detection · CNN · Deep learning · Accuracy

1 Introduction

The use of artificial intelligence systems [1–4] is increasingly advancing, especially in assisting clinical decision-making processes such as history taking, diagnosis, therapy, and prognosis. While such systems serve as aids to clinicians, they are particularly crucial for novice clinicians. To date, there have been several achievements in line with individuals' expertise, such as in medicine, education, or agriculture. In the field of healthcare, there has been a shift from manual disease analysis to more efficient and labor-saving disease analysis using expert systems or tools. Expert systems are an application of artificial intelligence techniques that mimic what an expert does when dealing with complex problems based on their knowledge.

Pneumonia is a leading cause of death among children under five years old (toddlers) worldwide, surpassing other diseases like AIDS, Malaria, and Measles. However, it has not received as much attention. Globally, more than 2 million toddlers die each year due to pneumonia, which equates to around 4 toddlers dying every minute. One out of every five toddler deaths is caused by pneumonia. In Indonesia, the national prevalence of Acute Respiratory Infections (ARI) [5–7] is 25.5% (16 provinces above the national average), and morbidity rates for pneumonia in infants and toddlers are 2.2% and 3%, respectively. Mortality rates for infants and toddlers are 23.8% and 15.5%, respectively. Pneumonia is one of the leading causes of child mortality [8] worldwide. The World Health Organization (WHO) estimates that it accounts for 15% of all deaths in children under the age of 5. In 2015, more than 900,000 children died due to pneumonia. In Indonesia alone, it is estimated that pneumonia claimed the lives of approximately 25,000 toddlers in 2013. Pneumonia can vary in severity, from mild cases to life-threatening ones, and can be caused by various organisms, including bacteria or viruses.

One of the diagnostic methods for pneumonia is through X-rays [9–12]. This diagnostic method provides an image of the lung conditions affected by pneumonia. The diagnosis is then analyzed by medical experts to determine the presence of pneumonia or not. X-ray or radiography is one of the medical imaging techniques [13–15] that use electromagnetic radiation to capture images or pictures of the interior of the body. X-ray procedures are quick and painless. This procedure is part of diagnostic support for more accurate diagnosis [16, 17] (see Table 1).

Considering the above-mentioned issues, the author is interested in conducting research titled "Improving the accuracy of pneumonia identification in children under 5 years old using the convolutional neural network method." The hope is that this research will provide problem-solving solutions and also help expedite the diagnosis process. The research uses methods for detecting and recognizing objects in an image [23–26] using artificial neural networks, one of which is the Convolutional Neural Network (CNN) algorithm, which is commonly used in image data processing [8]. The Convolutional Neural Network (CNN) [27] is one of the methods within

Table 1 Literature study

Title	Year	Author	Method	Result
Convolutional Neural Network for Automatic Pneumonia Detection in Chest Radiography [18]	2020	Septy Aminatul Khoiriyah, Arif Basofi, Arna Fariza	3rd CNN layers	The experimental results indicate that the augmentation strategy applied to the proposed CNN architecture resulted in an accuracy rate of 83.38%, whereas without augmentation, the accuracy rate was 80.25%
Classification of Pneumonia and Covid-19 Based On X-ray Image Using The Deep Residual Network Architecture [19]	2022	Prayogi Pangersa Illahi, Hilman Fauzi, Thomhert Suprapto Siadari	Using a sample of 233 X-ray images and employing the Adam optimizer	The experimental results show that the proposed CNN architecture achieved an accuracy of 87.7% with a loss value of 1.8%
CNN-based Deep Learning Model for Chest X-ray Health Classification Using TensorFlow [20]	2020	Rogelio Ruzcko (NMI) Tobias, Luigi Carlo M. De Jesus, Matt Ervin G. Mital, Sandy C. Lauguico, Marielet A. Guillermo, Edwin Sybingco, Argel A. Bandala, Elmer P. Dadios	Training, validation, and cross-entropy were performed using VGG32 as the pretrained model and functioning as a general convolutional neural network system	Accurate results were consistently achieved within a percentage range of over 90% during testing The confusion matrix provides numerical values for true positives (true normal), true negatives (true pneumonia), false positives (false normal), and false negatives (false pneumonia) data

(continued)

Table 1 (continued)

Title	Year	Author	Method	Result
Automated Pneumonia Diagnosis using a Customized Sequential Convolutional Neural Network [21]	2019	Raheel Siddiq	The 18-layer CNN model uses the Adam optimizer with 234 samples of normal and 340 samples of pneumonia	The classification model for 'normal' and 'Pneumonia' achieved a classification accuracy of 94.3%. The model also exhibited high sensitivity, reaching 99%
X-ray Chest Image Classification by A Small Sized Convolutional Neural Network [22]	2019	Ege Kesim, Zumray Dokur, Tamer Olmez	Using smaller image sizes with dimensions of 128 × 128 pixels	The CNN's performance during training is quite satisfactory, but there are some challenges. The use of different image files during the training process reduces the network's generalization ability. Additionally, due to its large size, it is not suitable for real-time applications. In this study, a new network structure and image size were processed for X-ray classification. Twelve classes of X-rays were obtained with a success rate of around 86% using the proposed network

deep learning that is widely used to address issues related to object detection and image classification [28–30]. Object detection has recently become one of the most interesting areas in computer vision and artificial intelligence (AI).

Object detection is a computer technology related to computer vision and image processing that involves detecting an object in digital images, which can include color and shape. The Convolutional Neural Network (CNN) is widely used in earlier research because it has a relatively high level of accuracy and has shown significant results in image recognition. One of the software that can be utilized for this purpose is the Python programming language, which falls under artificial intelligence software used for detecting various objects based on image datasets. There are several contributions from this research:

1. **Application of Expert Systems**: It discusses the transition from manual disease analysis to more efficient and labor-saving disease analysis using expert systems or tools. This reflects the application of artificial intelligence techniques to mimic the decision-making processes of experts when dealing with complex medical problems.
2. **Addressing Pneumonia Mortality**: The article sheds light on the critical issue of pneumonia-related child mortality, emphasizing that pneumonia is a major cause of death among children under five years old worldwide. It draws attention to the urgency of addressing this problem.

3. **Prevalence in Indonesia**: The article provides specific data on the prevalence of pneumonia in Indonesia, highlighting the significant impact of the disease in the country. It discusses not only the morbidity rates but also the high mortality rates among infants and toddlers.
4. **Potential Solutions**: The research project titled "Improving the accuracy of pneumonia identification in children under 5 years old using the convolutional neural network method" offers a potential solution to the challenges associated with pneumonia diagnosis. It proposes the use of artificial neural networks, specifically Convolutional Neural Networks (CNNs), to enhance the accuracy and speed of diagnosis, particularly in young children.
5. **Practical Implementation**: The research project's practical implementation of AI and CNN techniques in diagnosing pneumonia based on medical images demonstrates a real-world application of artificial intelligence in healthcare, potentially improving the efficiency and accuracy of diagnosis.

Overall, this research contributes to the understanding of the role of artificial intelligence in healthcare, specifically in addressing the urgent issue of pneumonia-related child mortality, and it proposes a potential solution to enhance pneumonia diagnosis using advanced AI techniques.

2 Materials and Methods

2.1 Dataset

The dataset used in this research consists of X-Ray scan images. The data used in this study is secondary data (public data) obtained from Kaggle, specifically the Chest X-ray images (Pneumonia) dataset, which comprises a total of 5,863 images. This dataset is categorized into two main classes: pneumonia and normal. The data was collected from a group of pediatric patients aged one to five years at Guangzhou Women and Children's Medical Center in Guangzhou, China. Using this dataset, the research aims to identify and analyze cases of pneumonia in children within this age range. The research methodology involves various stages, including a literature review to understand the medical context, experimental testing to analyze the X-ray images, and other relevant steps. Through the combination of these methods, this research seeks to provide a better understanding of pneumonia cases in children aged one to five years, with the hope of aiding in the improvement of diagnosis and management of this disease.

2.2 *Proposed Method*

Figure 1 shown the proposed method in this research involves several critical steps that are essential for the successful execution of the study. Each of these steps contributes to the overall methodology and helps ensure the accuracy and reliability of the findings. Here's a more detailed argumentation for each step:

1. **Data Collection**: The initial phase of this research involves the collection of data. In this case, the data is obtained from publicly available sources on Kaggle, specifically the Chest X-ray images (Pneumonia) dataset. This dataset is a valuable resource for studying pneumonia cases in X-Ray images. The careful selection of a reputable source for the data is crucial to ensure the quality and authenticity of the information.
2. **Data Preprocessing**: After acquiring the dataset, the next step is data preprocessing. This phase encompasses various tasks such as data cleaning, normalization, and augmentation. Cleaning involves the removal of any noisy or irrelevant data points, ensuring that the dataset is of high quality. Normalization is essential to standardize the data and make it suitable for deep learning algorithms. Data augmentation techniques, such as image resizing and cropping, may also be applied to enhance the dataset's diversity and robustness.
3. **Deep Learning Methodology**: Deep learning is at the core of this research. It involves the utilization of sophisticated neural network architectures to automatically learn and extract features from the X-Ray images. Convolutional Neural Networks (CNNs), for instance, are commonly used in image classification tasks due to their ability to capture intricate patterns within the data. The choice

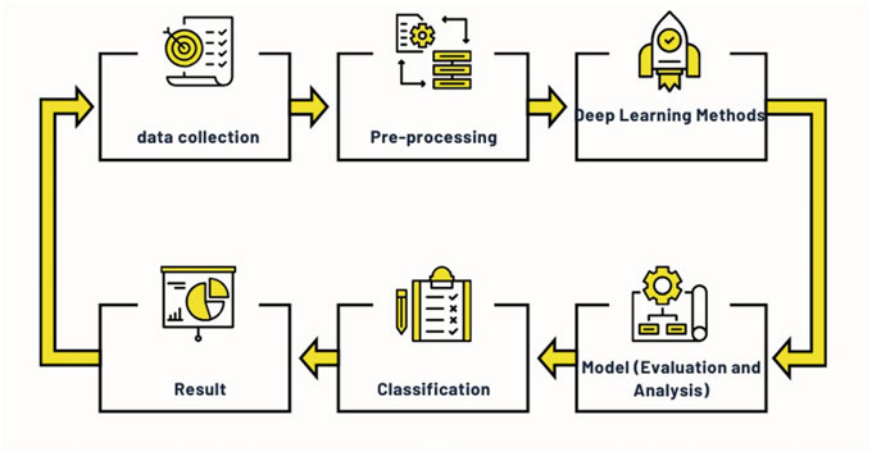


Fig. 1 Proposes method to enchance accuracy in the detection of pneumonia using convolutional neural networks

of the deep learning method is critical, and it should be justified based on its effectiveness in handling medical image data.

4. **Modeling (Evaluation and Analysis)**: Developing the deep learning models is a crucial step. This phase includes tasks such as model design, training, and evaluation. Researchers should carefully design the architecture of the neural network, select appropriate hyperparameters, and validate the model's performance. Evaluation metrics, such as accuracy, precision, recall, and F1-score, should be used to assess the model's effectiveness in classifying X-Ray images into the 'pneumonia' and 'normal' categories. The analysis of model performance and results should be thorough and supported by statistical evidence.
5. **Deployment**: Once a robust deep learning model is developed and evaluated, the next step is deployment. This involves making the model accessible and usable in real-world applications. Deployment considerations may include creating a user-friendly interface for healthcare professionals to upload X-Ray images and receive automated classifications, ensuring model scalability, and addressing security and privacy concerns.
6. **Classifying the Picture and Getting the Result**: In the deployment phase, X-Ray images are classified using the trained deep learning model. The results obtained from this classification process are crucial, as they determine the accuracy and reliability of the model in diagnosing pneumonia cases. Clear documentation of the classification results is essential, including the details of true positives, true negatives, false positives, and false negatives.
7. **Conclusion and Finishing**: In the final phase of the research, a comprehensive conclusion is drawn based on the analysis of results. Researchers should discuss the implications of their findings, including the potential impact on diagnosing pneumonia in pediatric patients aged one to five years. Any limitations of the study should be acknowledged, and future research directions can be proposed. Proper documentation and reporting are essential to ensure the research can be replicated and validated by other experts in the field.

The research process involves a systematic and well-structured approach, from data collection and preprocessing to deep learning modeling, deployment, and analysis of results. Each step plays a vital role in achieving the research objectives and contributing to our understanding of pneumonia diagnosis in young children.

2.3 Data Processing

Figure 2 shown that the data processing explains how and from where the data for this research is sourced, including primary and secondary data. Primary data refers to raw data collected by the researcher themselves (not by others) from primary sources for the purpose of their research, and such data did not previously exist. The data used in this research is secondary data (public data) obtained from Kaggle, specifically the Chest X-Ray Images (Pneumonia) dataset, consisting of 5,863 images categorized

into two groups: pneumonia and normal. This dataset was sourced from a group of pediatric patients aged one to five years at the Guangzhou Women and Children's Medical Center in Guangzhou, China. The research methods employed include literature review, experimental testing, and others. The application of deep learning in this research employs the Convolutional Neural Network (CNN) method. CNN is chosen because it has high recognition capabilities through deep learning processes. Within the CNN method, various architectures have been developed through experiments conducted by previous researchers. Convolutional layers use filters to extract objects from input images. These filters contain weights used to detect features of objects such as edges, curves, or colors. Convolution generates a linear transformation of the input image that corresponds to spatial information in the data. Filters are applied repeatedly, resulting in a series of receptive fields. There are parameters that can be adjusted to modify the characteristics of each layer, including filter size, stride, and padding. Stride controls how the filter is applied to the input data by moving along a predetermined pixel size. In summary, the research involves the use of secondary data from Kaggle's Chest X-Ray Images (Pneumonia) dataset, and it employs the CNN method for deep learning. The CNN method is chosen due to its effective feature extraction capabilities, and various architectures within CNN have been explored through prior experimentation.

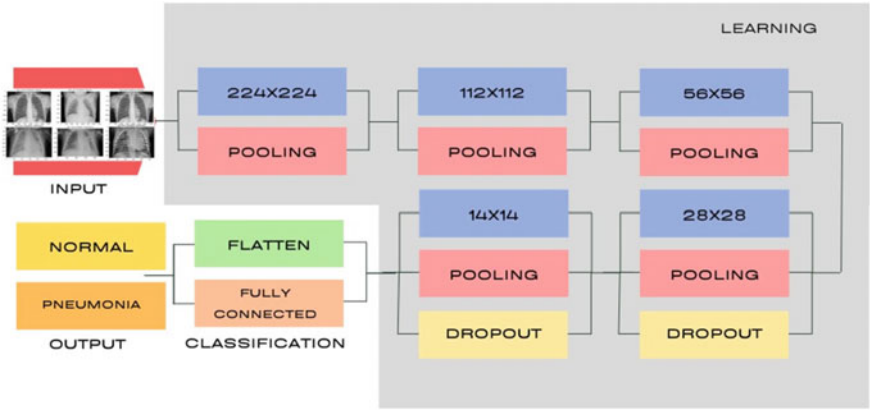


Fig. 2 Data processing to enhance accuracy in the detection of pneumonia using convolutional neural networks

3 Result

3.1 Training Process

Figure 3 involves the evaluation of the research results regarding the model used during the experiments, which is the Convolutional Neural Network (CNN). The evaluation of these results utilizes accuracy, recall, and precision values. After the model training is completed, the progression of the loss values on the test data is observed from the beginning of the training until the end of the training to identify when the model starts to overfit and when it reaches the minimum loss value on the test data. The CNN model with the minimum loss value on the test data is selected and then evaluated to assess the performance of each model in predicting each data point. The model, which has been trained with the training data, will produce a model ready for evaluation. Evaluation is performed by examining the confusion matrix, a classification matrix that includes the precision and recall of each class, as well as the overall accuracy of the model. This is done to determine how well the selected model performs in classifying data in each class. Additionally, a time plot is created to visualize the duration of time required during the model training. Subsequently, classification is performed based on the predetermined classes to facilitate predictions for normal patients and patients with pneumonia.

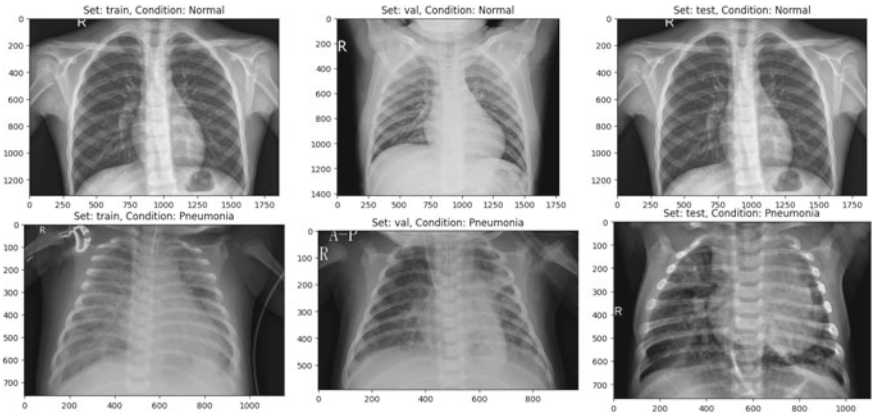


Fig. 3 Data training to enhance accuracy in the detection of pneumonia using convolutional neural networks

3.2 Training Process Results

During this phase, tests are conducted to evaluate the compatibility between the data and the suggested algorithm model. The selection of the model and the architecture employed are crucial in ensuring that the chosen model is appropriate for testing the dataset. After completing the hyperparameter tuning phase, the model will undergo training using picture data. The optimization technique employed in this study was Adam, facilitating the training process on individual subsets of data comprising 5722 training instances. Initially, the input data, which has dimensions of 224×224 pixels, comprises a randomly selected subset of 5722 data points from the training data set. Each data point is sent through the first hidden layer, composed of multiple components, notably convolution, activation, and pooling. The convolution component performs a convolution operation with a kernel size of 4×4 and a stride of 2. This procedure yields an output consisting of 56 feature maps, each measuring 56×56. The complete activation of the feature map is achieved by applying the ReLU activation function to each individual value inside the feature map. Following the activation process, the last step involves entering the pooling portion.

This study's technique is maximum pooling, utilizing a kernel size of 2×2 and a stride of 2. After traversing the initial layer, the output from said layer will proceed to the subsequent hidden layer. The operations performed in the second hidden layer are identical to those preceding. This pattern continues for the third, fourth, and fifth layers. The feature map generated from hidden layer 5 is propagated to the classification layer. Subsequently, the value produced by the output layer of batch data is utilized to compute the overall loss value derived from the loss function. The loss function employed in this study is cross-entropy loss. The progression from the input layer to the computation of the output layer is referred to as the forward phase. Following the forward phase, the CNN model undergoes a reverse process to compute the gradient for each parameter, specifically the weight and bias, which need to be optimized. Subsequently, the aforementioned parameters undergo an update using newly derived parameters through the Adam optimization method. This particular stage is commonly referred to as the backward phase. The comprehensive sequence encompassing both the forward and backward phases is called backpropagation. One epoch is a single iteration of the backpropagation procedure applied to all data segments. The research methodology involves doing the process for five epochs, during which the model with the lowest loss value in the test data will be preserved.

This study will conduct two rounds of training for the CNN model. The first round will involve under sampled data, while the second round will involve oversampled data. To create training loss values and training accuracy values, it is necessary to do appropriate training. The findings for training loss were observed to be 0.2, as depicted in Fig. 4. Additionally, the training accuracy was determined to be 0.96 or 96%.

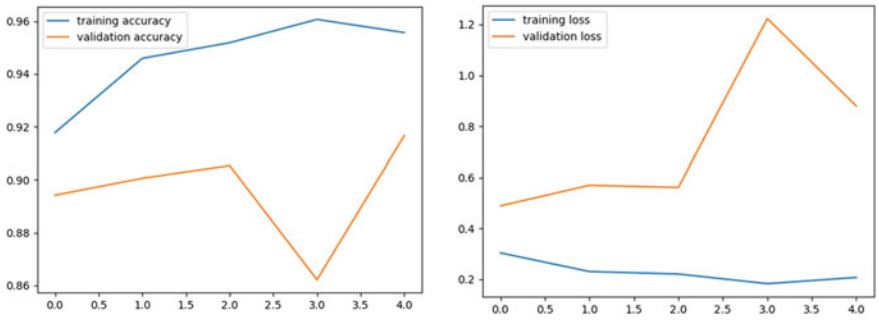


Fig. 4 Training loss and training accuracy graphs

3.3 Validation Process Results

The constructed convolutional neural network (CNN) model comprises five distinct components. The Convolutional Neural Network (CNN) architecture consists of layers, where each block and the fully linked layer are constructed using an input image size of 224 × 224 pixels. The dimensions of the kernel utilized in the convolution process are 4 × 4.

The validation method assesses the degree of recognition accuracy by reading data. The greater the amount of data used for training, the higher the level of accuracy achieved. The accuracy value of each epoch can be observed. Table 2 presents a comparative analysis of Accuracy and Loss at Each Epoch.

The correctness of the model testing findings is determined by computing the confusion matrix on the chest X-ray image. There were a total of 1342 chests of normal size. In a dataset consisting of 3876 pneumonia chest X-ray images, it was seen that 3788 of these images were accurately classified as pneumonia cases, referred to as True Pneumonia (TP). Additionally, 14 chest X-ray photos were misclassified as pneumonia cases.

$$\text{Accuracy} = \frac{TN + TP}{TN + TP + FN + FP} \times 100\% \tag{1}$$

$$\text{Presisi} = \frac{TN}{TN + FN} \times 100\% \tag{2}$$

Table 2 Comparison of accuracy and loss at each epoch

Epoch		1	2	3	4	5
Training	Loss	0,3037	0,2304	0,2205	0,1826	0,2070
	Accuracy	0,9178	0,9459	0,9519	0,9607	0,9557
Validation	Loss	0,4886	0,5695	0,5611	1,2234	0,8790
	Accuracy	0,8	0,90	0,90	0,86	0,90

$$\text{Recall} = \frac{TN}{TN + FP} \times 100\% \tag{3}$$

where TP is True Positives (true positives), TN is True Negatives (true negatives), FP is False Positives (false positives), and FN is False Negatives (false negatives).

3.4 Enhancing Accuracy in the Detection of Pneumonia

Difference in number of images

In this study, the researchers attempted to conduct experiments utilizing datasets with varying proportions relative to the standard distribution of normal and pneumonia data. Specifically, the dataset proportions examined were 300:700, 800:1200, 900:2100, 1000:3000, and 1800:3200.

According to the statistics presented in Table 3, it is evident that Experiment 5 exhibits the highest validation accuracy (0.9167). This finding implies that Experiment 5 performs better in accurately identifying unknown data samples. Nevertheless, it should be noted that this particular model exhibits the lengthiest training duration, amounting to 11 h and 41 min. This extended training time may present certain challenges in certain practical situations. If one prioritizes achieving validation accuracy and possesses sufficient processing resources and time, Experiment 5 might be regarded as the most optimal dataset. Nevertheless, when a trade-off between accuracy and training time is required, Experiment 2 presents a satisfactory validation accuracy of 0.8915, coupled with a comparatively shorter training duration of 3 h and 53 min.

Table 3 Test results with differences in the number of images

No.	Epoch	Number	Size	Accuracy validation	Training accuracy	Time
1	5	Normal 300 pneumonia 700	224 × 224	0.8771	0.9504	2 h 41 min
2	5	Normal 800 Pneumonia 1200	224 × 224	0.8915	0.9658	3 h 53 min
3	5	Normal 900 Pneumonia 2100	224 × 224	0.8997	0.9504	7 h 16 min
4	5	Normal 1000 Pneumonia 3000	224 × 224	0.9009	0.9612	8 h 39 min
5	5	Normal 1800 Pneumonia 3200	224 × 224	0.9167	0.9557	11 h 41 min

Difference in epoch value

In this study, the investigators conducted experiments by manipulating the epoch values. Specifically, the epochs utilized in the studies were 3, 5, 10, 15, and 20.

Table 4 shows that Experiment 5 exhibits the best validation accuracy (0.9721) and the highest training accuracy (0.9773), indicating a robust performance in both training and validation phases. Nevertheless, it is worth noting that this model exhibits the lengthiest duration for training, amounting to 28 h and 53 min. This extended training time may provide challenges in certain practical situations. If one emphasizes both validation and training accuracy and possesses sufficient computational resources and time, Experiment 5 might be regarded as the most optimal dataset. Nevertheless, when a trade-off between accuracy and training time is desired, it is worth noting that Experiment 4 exhibits a notable level of validation accuracy (0.9615) and training accuracy (0.9677) while requiring a comparatively lower training duration of 27 h and 09 min

Difference in size value

In this study, the investigators conducted experiments by manipulating the size. Specifically, the sizes utilized in the studies were 64 × 64, 128 × 128, and 244 × 244.

Table 5 shows that Experiment 3 exhibits the highest validation accuracy (0.9443) and a comparatively high training accuracy (0.9402). This observation suggests a high level of proficiency in both the training and validation processes. Furthermore, it is worth noting that Experiment 2 exhibits a commendable level of validation accuracy. However, Experiment 3 surpasses Experiment 2 in terms of validation accuracy, albeit by a little margin, and also demonstrates satisfactory performance in terms of training accuracy. Experiment 1 exhibits a notably diminished validation accuracy of 0.6254 in comparison to the remaining trials, raising doubts regarding its suitability as the optimal choice.

Table 4 Test results with differences in the number of epochs

No.	Epoch	Number	Size	Accuracy validation	Training accuracy	Time
1	3	Normal 800 Pneumonia 1200	224 × 224	0.8821	0.9523	3 h 54 min
2	5	Normal 800 Pneumonia 1200	224 × 224	0.8920	0.9471	4 h 37 min
3	10	Normal 800 Pneumonia 1200	224 × 224	0.9020	0.9571	25 h 20 min
4	15	Normal 800 Pneumonia 1200	224 × 224	0.9615	0.9677	27 h 09 min
5	20	Normal 800 Pneumonia 1200	224 × 224	0.9721	0.9773	28 h 53 min

Table 5 Test results with differences in size

No.	Epoch	Number	Size	Accuracy validation	Training accuracy	Time
1	5	Normal 800 Pneumonia 1200	64 × 64	0,6254	0,9207	1 h 32 min
2	5	Normal 800 Pneumonia 1200	128 × 128	0,9185	0,9535	4 h 2 min
3	5	Normal 800 Pneumonia 1200	244 × 244	0,9443	0,9402	5 h 41 min

4 Discussion

The interpretation of the research results provides valuable insights into the impact of various parameters on the performance of the Convolutional Neural Network (CNN) model for pneumonia detection. We compare the effects of different factors including processing time, dataset size, image dimensions, and hidden layers.

Comparison of Processing Time: The experiments demonstrated that as we increased the complexity of the CNN architecture by adding more hidden layers, the processing time also increased. While utilizing five hidden layers led to higher accuracy rates, it came at the cost of extended training durations. This indicates a trade-off between model complexity and computational resources. Researchers must carefully consider the balance between accuracy and practical feasibility when designing CNN architectures.

Comparison of Dataset Size: The substantial increase in dataset size from 233 images to 5,000 images significantly improved the model's ability to generalize and make accurate predictions. However, it also led to increased training time. This suggests that a larger dataset can enhance the model's performance, but researchers need to allocate sufficient computational resources to accommodate the expanded dataset.

Comparison of Image Sizes: Experimentation with different image sizes revealed that using 244 × 244 dimensions yielded the most optimal results. This is in contrast to 128 × 128, which did not produce the same level of accuracy. The choice of image size is crucial in preserving essential details while maintaining computational efficiency. Larger image sizes provide more information for the model to learn from, but they also demand greater computational resources.

Comparison of Hidden Layers: The transition from a three-layered architecture to five hidden layers marked a significant improvement in accuracy. However, it was accompanied by an increase in processing time. This highlights the importance of striking a balance between model complexity and practical implementation. A more intricate architecture may yield higher accuracy, but it may also require more computational power and time.

The Impact from this Research: Collectively demonstrate the pursuit of finding the most optimal approach for detecting pneumonia. It emphasizes the need for a careful consideration of parameters to achieve the right balance between accuracy and resource utilization. This research contributes valuable insights into the utilization of CNNs for medical image analysis and lays the groundwork for further advancements in pneumonia detection. The findings are poised to have a substantial impact on improving diagnostic outcomes, ultimately benefiting patient care in the field of respiratory health.

5 Conclusion

This paper presents a study that centers on the crucial matter of pneumonia identification in adult patients. It emphasizes the potential of modern artificial intelligence approaches, particularly Convolutional Neural Networks (CNNs), in enhancing diagnostic accuracy to a substantial extent. The study conducted a series of tests to systematically investigate the influence of different parameters on the performance of the Convolutional Neural Network (CNN) model.

The primary discoveries of this study suggest that getting ideal outcomes requires careful examination of criteria such as the number of photos, epoch values, and image size. In Experiment 5, a dataset including 1800 normal photos and 3200 pneumonia images was employed, resulting in the attainment of the greatest validation accuracy of 91%. Nevertheless, it is crucial to recognize that this particular experiment also entailed the most extensive duration for training, indicating a potential compromise between precision and computing capabilities.

The study's outcomes hold substantial promise for the field of medical imaging and diagnostic processes. By leveraging CNNs, the research provides a viable path toward more reliable and accurate pneumonia diagnoses in adult patients. This advancement is particularly significant in the context of respiratory conditions, where timely and precise diagnosis can significantly impact patient outcomes.

This research provides significant contributions to the field by examining the utilization of Convolutional Neural Networks (CNNs) to improve the identification of pneumonia in adult individuals. The results underscore the significance of parameter optimization and lay the groundwork for future investigations in this domain. The study's findings have substantial implications for enhancing patient care, diagnostic outcomes, and the entire field of medical imaging within the realm of respiratory health.

References

1. Talo M (2019) Pneumonia detection from radiography images using convolutional neural networks. In: 27th signal processing and communications applications conference, SIU 2019, pp 1–4. https://doi.org/10.1109/SIU.2019.8806614
2. Li X, Chen F, Hao H, Li M (2020) A pneumonia detection method based on improved convolutional neural network. In: Proceedings of the 2020 IEEE 4th information technology, networking, electronic and automation control conference. ITNEC 2020, no Itnec, pp 488–493. https://doi.org/10.1109/ITNEC48623.2020.9084734
3. Nayyar A, Jain R, Upadhyay Y (2020) Object detection based approach for Automatic detection of Pneumonia. In: International conference on artificial intelligence, big data, computing and data communication systems, icABCD 2020—Proceedings. https://doi.org/10.1109/icABCD49160.2020.9183876
4. Zhang J et al (2021) Viral pneumonia screening on chest X-rays using confidence-aware anomaly detection. IEEE Trans Med Imaging 40(3):879–890. https://doi.org/10.1109/TMI.2020.3040950
5. Subiyanto, Mulwinda A, Andriani D (2017) Intelligent diagnosis system for acute respiratory infection in infants. In: Proceeding—2017 3rd international conference on science in information technology theory application of IT for education, industry and society in big data era, ICSITech 2017, vol 2018-Janua, pp 558–562. https://doi.org/10.1109/ICSITech.2017.8257175
6. Ertugrul DC, Elci A, Bitirim Y (2017) An intelligent tracking system: application to acute respiratory tract infection (TrackARTI). In: Proceedings of the international conference on computers, software, and applications, vol 2, pp 137–142. https://doi.org/10.1109/COMPSAC.2017.58
7. Trandafir LM, Boiculese LV, Dimitriu G, Moscalu M (2017) Recurrent respiratory tract infections in children. In: 2017 E-Health and bioengineering conference, EHB 2017, pp 741–744. https://doi.org/10.1109/EHB.2017.7995530
8. Gabruseva T, Poplavskiy D, Kalinin A (2020) Deep learning for automatic pneumonia detection. In: IEEE conference on computer vision and pattern recognition workshops, vol 2020-June, pp 1436–1443. https://doi.org/10.1109/CVPRW50498.2020.00183
9. Chowdhury YS, Dasgupta R, Nanda S (2021) Analysis of various optimizer on CNN model in the application of pneumonia detection. In: 2021 3rd international conference on signal processing and communication, ICPSC 2021, no May, pp 417–421. https://doi.org/10.1109/ICSPC51351.2021.9451768
10. Cococi A, Felea I, Armanda D, Dogaru R (2020) Pneumonia' detection on Chest X-ray, images using convolutional neural networks' designed for 5 source & constrained environments, pp 31–34
11. Islam SR, Maity SP, Ray AK, Mandal M (2019) Automatic detection of pneumonia on compressed sensing images using deep learning. In: 2019 IEEE Canadian conference of electrical and computer engineering, CCECE 2019, pp 1–4. https://doi.org/10.1109/CCECE.2019.8861969
12. Sharma H, Jain JS, Bansal P, Gupta S (2020) Feature extraction and classification of chest X-ray images using CNN to detect pneumonia. In: Proceeding of the Confluence-2020:10th international conference on cloud computing, data science & engineering, pp 227–231. https://doi.org/10.1109/Confluence47617.2020.9057809
13. Kajla V, Gupta A, Khatak A (2018) Analysis of x-ray images with image processing techniques: a review. In: 2018 4th international conference on computing communication and automation, ICCCA 2018, pp 1–4. https://doi.org/10.1109/CCAA.2018.8777693
14. Tang J, Jiao X (2012) Research progress of digitized medical X-ray imaging system. In: Proceedings of the 2012 international conference on computer science and information processing, CSIP 2012, pp 1074–1077. https://doi.org/10.1109/CSIP.2012.6309042
15. Bulut S, Özçinar A, Ftç Çİ, Akpek A, Yeni İ (2015) X-Işını Görüntülerinde Segmentasyon ve Kırık Tespiti için Yeni Algoritma. A new algorithm for segmentation and fracture detection in X-ray images, pp 1–4

16. Daoud MI, Alrahahleh Y, Abdel-Rahman S, Alsaify BA, Alazrai R (2021) COVID-19 diagnosis in chest X-ray images by combining pre-trained CNN models with flat and hierarchical classification approaches. In: 2021 12th international conference on information and communication systems, ICICS 2021, pp 330–335. https://doi.org/10.1109/ICICS52457.2021.9464532
17. Chowdhury MEH et al (2020) Can AI help in screening viral and COVID-19 pneumonia? IEEE Access 8:132665–132676. https://doi.org/10.1109/ACCESS.2020.3010287
18. Khoiriyah SA, Basofi A, Fariza A (2020) Convolutional neural network for automatic pneumonia detection in chest radiography. In: IES 2020—international electronics symposium, pp 476–480. https://doi.org/10.1109/IES50839.2020.9231540
19. Illahi PP, Fauzi H, Siadari TS (2022) Klasifikasi penyakit pneumonia Dan Covid-19 berbasis citra X-ray menggunakan arsitektur deep residual network. eProc Eng 9(4):1837–1843
20. Tobias RRNMI et al (2020) CNN-based deep learning model for chest X-ray health classification using tensorflow. In: Proceedings of the 2020 RIVF international conference on computing & communication technologies, RIVF 2020. https://doi.org/10.1109/RIVF48685.2020.9140733
21. Siddiqi R (2019) Automated pneumonia diagnosis using a customized sequential convolutional neural network. In: ACM international conference proceeding series, pp 64–70. https://doi.org/10.1145/3342999.3343001
22. Kesim E, Dokur Z, Olmez T (2019) X-ray chest image classification by a small-sized convolutional neural network. In: 2019 scientific meeting on electrical-electronics & biomedical engineering and computer science, EBBT 2019, pp 1–5. https://doi.org/10.1109/EBBT.2019.8742050
23. Akshaya, Ranjan Kumar HS, Bhat D (2019) A novel method to recognize object in images using convolution neural networks. In: 2019 international conference on intelligent computing and control systems, ICCS 2019, no Iciccs, pp 425–430. https://doi.org/10.1109/ICCS45141.2019.9065367
24. Abegg F, Worn H (2000) Robust algorithms for recognizing shape changes, pp 335–338
25. Feryanto A, Supriana I (2011) Location recognition using detected objects in an image. In: Proceedings of the 2011 international conference on electrical engineering and informatics, ICEEI 2011, no July. https://doi.org/10.1109/ICEEI.2011.6021589
26. Yu L, Chen X, Zhou S (2018) Research of image main objects detection algorithm based on deep learning. In: 2018 3rd IEEE international conference on image, vision and computing, ICIVC 2018, pp 70–75. https://doi.org/10.1109/ICIVC.2018.8492803
27. Labhane G, Pansare R, Maheshwari S, Tiwari R, Shukla A (2020) Detection of pediatric pneumonia from chest X-ray images using CNN and transfer learning. In: Proceedings of the 3rd international conference on emerging technologies in computer engineering: machine learning and internet of things, ICETCE 2020, no February, pp 85–92. https://doi.org/10.1109/ICETCE48199.2020.9091755
28. Pradhan PK, Baruah U (2023) Advertisement image classification using convolution neural network. Lecture notes in networks and systems, vol 498, pp 225–242. https://doi.org/10.1007/978-981-19-5090-2_22
29. Gao Z, Edirisinghe E, Chesnokov S (2019) Image super-resolution using CNN optimised by self-feature loss. Eran Edirisinghe Department of Computer Science, Loughborough University, UK Imaging & Vision Group, Arm Ltd., UK. In: 2019 IEEE international conference on image processing, pp 2816–2820
30. Kido S, Hirano Y, Hashimoto N (2018) Detection and classification of lung abnormalities by use of convolutional neural network (CNN) and regions with CNN features (R-CNN). In: 2018 international workshop on advanced image technology, IWAIT 2018, pp 1–4. https://doi.org/10.1109/IWAIT.2018.8369798

Development of a Baby Incubator with a Vital Sign Monitoring Tool Based on Pan-Tompkins Method for Heart Rate Calculation



Bambang Guruh Irianto, Anita Miftahul Maghfiroh, M. Ikhsan Firmansyah, Abd. Kholiq, Syevana Dita Musvika, and Yuni Kusmiyati

Abstract Heart signals and heart rate are important aspects that need to be monitored in the care of premature babies because they describe the health condition of premature babies. This research aims to develop a baby incubator with accurate Electrocardiogram (ECG) monitoring that can facilitate monitoring of the baby's condition directly so as to get the right treatment. This research uses an Electrocardiograph circuit, temperature sensor, NTC skin sensor, loadcell sensor, and ultrasonic sensor. Arduino is used as a processor of all data and in ECG processing the Pan-Tompkins algorithm is used to process the QRS signal which will then detect the R peak and calculate the accurate heart rate value. The testing procedure of this module is carried out by taking data directly from 10 subjects and then the data obtained is processed to obtain accuracy regarding heart signals and heart rate. The main result of this research is the development of a baby incubator that can monitor heart signals in real time using the Pan-Tompkins algorithm. Measurements using a comparison tool showed an average heart rate measurement error of 0.2%. It was concluded that the incubator performance showed good results. It was concluded that the incubator performance showed good results. The results of this study can be used for doctors to monitor the condition of premature babies remotely using a smartphone or computer so that it is expected to accelerate the analysis and treatment of patients.

Keywords Baby incubator · ECG · Pan-Tompkins · Telemedicine · Heart Rate

B. G. Irianto · A. M. Maghfiroh (✉) · M. I. Firmansyah · Abd. Kholiq · S. D. Musvika
Department of Electromedical Engineering, Health Polytechnic Ministry of Health Surabaya, Surabaya, Indonesia
e-mail: anitamiftah@poltekkesdepkes-sby.ac.id

Y. Kusmiyati
Health Polytechnic Ministry of Health Yogyakarta, Yogyakarta, Indonesia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_42

1 Introduction

A baby incubator is a newborn warming aid and is often used for premature babies. It keeps the baby's body temperature and humidity stable, and it can help avoid respiratory infections and isolate neonates, particularly those born spontaneously [1–3]. Premature A premature baby is a baby born in abnormal conditions where the baby weighs less than 1500 g and the baby is smaller than average or the gestational age is less than 37 weeks [4, 5]. Monitoring the baby's health is very important, especially in detecting health problems in premature babies. Premature babies are cared for in intensive care units to ensure the baby's temperature and humidity match that of the mother's womb [6, 7]. Due to prematurity, the baby's physical conditions such as lungs, skin, respiratory system, and digestive system are not fully developed so it is difficult to adapt to life outside the womb, so the condition of premature babies must be continuously monitored and detected [8]. Premature infants with low birth weight are prone to hypothermia which is considered a leading cause of infant mortality. According to WHO, hypothermia for body temperature below the normal range (36.5–37.5 °C), includes low-temperature stress: 36.0–36.4 °C, moderate hypothermia: 32.0–35.9 °C, severe hypothermia: <32.0 °C [9–11]. According to the WHO, preventing hypothermia is recognized as an important part of care for all newborns. Up to 42% of neonatal deaths can be prevented through hypothermia prevention [12]. may cause apnea. Apnea is a condition where the baby stops breathing for a few seconds. One of the signs of apnea is a decrease in heart rate during ECG signal recording followed by a decrease in oxygen uptake. The number of heartbeats per minute, often known as beats per minute (BPM), is the heart rate. Medical experts utilize heart rate measures to help diagnose and monitor health conditions [13, 14]. The normal BPM value for infants is 120–150 BPM [15]. So it is very important to monitor ECG parameters and temperature, when symptoms like the above occur, action must be taken quickly, the baby's condition can deteriorate anytime and anywhere. Of course, a personal nurse/doctor is not always with the baby in the neonate room. For this reason, it is necessary to have a system that notifies the results of BPM and body temperature monitoring to the doctor to evaluate the baby's development and take action when the baby's heart rate or body temperature is abnormal [16, 17]. The heart is an important organ that circulates blood throughout the body. Knowing about the heart is essential for understanding a person's overall health. Heart disease is the largest cause of sudden mortality worldwide, followed by heart attacks. As a result, knowledge regarding a person's heart health is critical [18]. An electrocardiogram (ECG) can describe the electrical activity of the heart and is an important tool in the diagnosis of cardiovascular disease (CAD). With the rapid development of portable and wireless ECG technology, real-time and routine ECG monitoring is becoming increasingly important for medical care especially for the elderly [19]. Several articles and research works have been developed in this field and vary among several themes, including ECG signal filtering, extraction of QRS complexes and other waves from ECG signals, ECG signal compression, ECG

signal automatic processing methods, and live monitoring systems for heart conditions [20]. The Pan-Tompkins algorithm is a widely used tool to perform non-invasive ECG feature extraction, to detect physical conditions, as well as heart diseases. Since most of the energy associated with the ECG signal is located in the QRS complex, the detection of this part of the ECG signal is crucial for evaluating the patient's health condition. The QRS complex is the most striking waveform in the ECG signal, it serves as the basis for automatic determination of the heart rate, as well as the benchmark point for classifying the cardiac cycle and identifying any abnormalities. Over the past few decades, QRS complex detection has been extensively studied [21].

In 2018 Xuan Lu Yu et al. conducted research on QRS detection based on adaptive threshold enhancement in this study several algorithm steps were used including preprocessing, peak search, and adaptive threshold QRS detection. The detection rate value is 99.41%, sensitivity (Se) 99.72%, and specificity (Sp) 99.69%, these values are only obtained from data derived from the MIT-BIH Arrhythmia database [22]. Another study was conducted by Feifei Liu et al. on the accuracy of QRS detection based on the pan Tompkins method in general with a low-quality electrocardiogram database, this study was conducted by trying 3 pan Tompkins methods, namely the original method, the average-based modification method, modification method based on the middle value tested on 200 ecg data from physionet (100 good quality database and 100 low quality data) obtained the results that in high quality data all pan Tompkins methods have high accuracy (>99%) and in low quality data the pan Tompkins modification method based on the middle value has the highest accuracy at a value of 76.03%, just like before this value was only obtained from data derived from the physionet database or MIT-BIH Arrhythmia [23]. Another study was conducted by Xueyu Wu et al. This study uses an optimized Pan-Tompkins algorithm to detect heartbeats based on the addition of P waves and the location of T waves. just like previous researchers, this algorithm was tested on 109,966 samples from the MIT-BIH arrhythmia database including various arrhythmias (such as atrioventricular block, tachycardia, early ventricular contractions, atrial flutter, premature heart rate, ventricular fibrillation) [24]. Other researchers, Amrita Rana, and Kyung Ki Kim, also conducted experiments on 8 patient ECG data from MIT-BIH using a modified Pan Tompkins algorithm to read conditions and abnormalities in the heart. This experiment was only simulated on MATLAB 2016a software with 90% accuracy results for two types of heart disease [25].

In some previous studies, the proposed pan Tompkins method was tested using data from MIT-BIH Arrhythmia and simulated through programs such as MATLAB but has not been used in real-time programs, while in Bambang Guruh Irianto et al. research, he designed a baby incubator equipped with ECG vital sign monitoring but still uses a fixed threshold where the BPM value obtained is still less accurate so that it can still be maximized again in the BPM calculation [26]. Based on some of the research above, we intend to develop a telemedicine-based baby incubator by applying the use of an adaptive threshold for R peak detection in the QRS signal owned by the pan-tompkins algorithm on vital sign monitoring of baby incubator devices which are then processed to obtain accurate heart rate values and can be displayed to a computer or android in real-time. This research aims to develop a

baby incubator with accurate ECG monitoring that can facilitate monitoring of the baby's condition directly so as to get the right treatment. The contribution of this paper is:

1. Accurate identification of R peaks in real-time ECG signals. The R-peak is an important marker in the ECG waveform because it represents the onset of ventricular depolarization. Accurate detection of R peaks is critical for a variety of clinical applications, including heart rate counting and arrhythmia detection.
2. Its efficient implementation ensures that R-peaks are detected promptly as ECG data is acquired. Real-time processing is essential for applications where immediate feedback or intervention is necessary, such as in critical care settings.
3. R-detection methods in the Pan-Tompkins algorithm are designed to be robust against noise and artifacts commonly present in ECG signals. Robustness against noise ensures that the algorithm can accurately detect R-peaks even in the presence of interference, leading to reliable results in various conditions.

2 Materials and Method

2.1 Experimental Setup

This study involved babies aged between 1 day and 3 months. Conducting research involving vulnerable populations such as infants requires careful ethical consideration and compliance with ethical guidelines and regulations. Our research has passed ethical tests and obtained an ethical approval certificate with certificate number 206/HRECC.FODM/II/2023 issued by Airlangga University Faculty of Dental Medicine Health Research Ethical Clearance Commission. This ethical approval ensures that the research has been reviewed and approved by an ethics committee or institutional review board to safeguard the rights, safety, and welfare of the participants. Informed consent to parents or legal guardians. The consent process outlines the nature of the study, potential risks and benefits, and how the baby data will be used. This study collects module data by comparing module readings and patient monitors. ECG signals used in the study were taken using an ECG circuit and processed using the pan-tompkins algorithm in the phyton program and will be displayed on both LCD and Android.

Materials and Tools. This study uses an ECG circuit In the system there is an ECG circuit made with the basic instrumentation circuit, the output of the basic instrumentation will be filtered to smooth the ECG signal, a filter for ECG signals with a lower limit frequency of 0.1 Hz, then the upper limit of 102 Hz and the output will be forwarded to the notch filter circuit which will reduce the voltage with a frequency of 50 Hz and only pass the voltage with a frequency other than 50 Hz. After that, go to the amplification circuit to do the amplification of the notch filter results. After that, it goes to the adder circuit which aims to increase the ECG signal reference so that all the signals are at a positive voltage so that they can be

read by the Arduino, the output of the ECG signal will be processed in the Arduino microcontroller. Then processed by Arduino sent to Raspberry and then processed using the pan-tompkins algorithm. starts from preprocessing data using a high pass filter with a transfer function based on Eq. (1) [24].

$$H(z) = \frac{-1 + 32z^{-16} + z^{-32}}{1 + z^{-1}} \tag{1}$$

The amplitude response is Eq. (2)

$$|\mathrm{H(wT)}| = \frac{\left[256 + \sin^2(16\omega \mathrm{T}\right]^{1/2}}{\cos(\mathrm{wT}/2)} \tag{2}$$

The difference equation is Eq. (3)

$$\mathrm{y(nT)} = 32 \times (\mathrm{nT} - 16\mathrm{T}) - \left[\mathrm{y(nT} - \mathrm{T}) + \mathrm{x(nT)} - \mathrm{x(nT} - 32\mathrm{T})\right] \tag{3}$$

where T is the sampling period, This filter's low cutoff frequency is approximately 5 Hz, its gain is 32, and its delay is 16 samples. Then a low pass filter with a transfer function based on Eq. (4)

$$H(z) = \frac{1 - 2z^{-6} + z^{-12}}{1 - 2z^{-1} + z^{-2}} \tag{4}$$

After filtering, the next step is a derivative filter to obtain slope information from the qrs signal using a transfer function based on Eq. (5).

$$H(z) = (-z^2 - 2z^1 + 2z^1 + z^2)\frac{1}{8T} \tag{5}$$

After several filtering processes, the process is carried out using the squaring function and continued with the moving window integration process based on Eq. (6)

$$\begin{aligned} Y(nT) &= (x(nT))^2 \\ Y(nT) &= \frac{(x(nT - (N-1)T) + x(nT - (N-2)T + \ldots x(nT))}{N} \end{aligned} \tag{6}$$

where T is the sampling period, then the last Fig is the threshold process using adaptive threshold so that the peak R of the QRS signal is obtained and the heart rate value can be calculated with Eq. (7)

$$HR = \frac{sampling\,rate \times 60}{R - Rinterval} \tag{7}$$

where HR is Heat rate. Test In this study, we tested the module by comparing the measurement results using the module designed with the measurement results on the patient monitor device. the ECG signal used in this test is lead II, where the subject of ECG signal collection used is babies aged between 1 day and 3 months.

2.2 Block Diagrams

Figure 1 is a block diagram of the module design used. In the block diagram there are three parts, namely: input, process, and output. In the initial condition input section, the signal is tapped from the electrodes that have been installed on the respondent using ECG lead II. The ECG data will be processed analogically by several circuits such as amplifiers, bandpass filters, notch filters to produce the required ECG signal and reduce noise or signals that are not needed. Then, the results of the ECG will enter the microcontroller and raspberry then will be processed using the pan-tompkins algorithm in phyton program then obtained QRS signal which is then calculated as heart rate value through an R-R interval.

From Fig. 1 it is explained that Electrodes are attached to the respondent to capture the ECG signal. The ECG Lead II configuration is often used, which involves placing one electrode on the right arm and another on the left leg, with the recording electrode on the left arm. The captured ECG data goes through analog processing circuits. These circuits can include Amplifiers: These increase the amplitude of the weak ECG signal for better processing. Bandpass Filters: These allow only the frequency range of the ECG signal (typically around 0.05–150 Hz) to pass through, filtering out unwanted noise and interference. Notch Filters: These are used to remove specific unwanted frequencies, such as powerline interference (50 or 60 Hz). The processed ECG signal is then sent to a microcontroller and a Raspberry Pi for further digital processing.

Digital Processing: On the microcontroller and Raspberry Pi, digital signal processing algorithms are employed. Specifically, you mentioned the "pan-tompkins

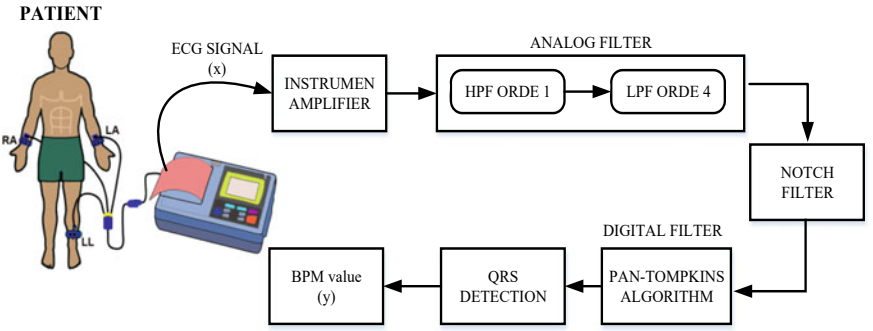


Fig. 1 Block diagram of the vital sign ECG on baby incubator

algorithm": This is a widely used algorithm for QRS complex detection in ECG signals. It's used to identify the QRS complex, which corresponds to the depolarization of the ventricles in the heart. The Pan-Tompkins algorithm is implemented in a Python program running on the Raspberry Pi. The output of this algorithm is the detected QRS complex locations in the ECG signal. Output Section: The QRS complex locations obtained from the Pan-Tompkins algorithm represent the R-wave peaks in the ECG signal. By calculating the time intervals between consecutive R-wave peaks (R-R intervals), the heart rate can be determined. The heart rate value is calculated based on the reciprocal of the average R-R interval. In summary, this module takes raw ECG data from electrodes attached to a respondent, processes it using analog circuits and digital processing algorithms, and finally calculates the heart rate based on the detected R-R intervals. It's a common setup for monitoring and analyzing heart activity in various medical and research applications.

3 Result

The measurement of this module is carried out by comparing the measurement results of the module with the results shown by the patient monitor, as a reference and comparison. Before data collection, a test was carried out on the Pan-Tompkins algorithm which will be used using data from Physionet. After that, data collection was carried out on the subject. This measurement used 10 adult research subjects where each research subject had criteria over 1 day and 3 months of age and data collection was carried out 5 times under the same conditions for each subject. The research in Fig. 2 is the result of testing the pantomkin method with Fig. 2a which is a band pass filter process Fig. 2b Derivative signal output Fig. 2c Squaring signal output Fig. 2d Moving Window Integrated signal output Fig. 2e and R Peak detection, this process can reduce noise to facilitate further processing.

Figure 2—Band Pass Filter Process: The initial step involves applying a bandpass filter to the input signal. A bandpass filter allows a specific range of frequencies to pass through while attenuating frequencies outside that range. This step likely helps in isolating the relevant frequency components of the signal while reducing noise. Figure 2b—Derivative Signal Output: In this Fig, the derivative of the signal obtained from the previous step is calculated. The derivative provides information about the rate of change of the signal. This step could be used to identify points of rapid change or transitions in the signal, which might correspond to certain events or features. Figure 3—Squaring Signal Output: Squaring the signal involves multiplying each data point by itself. This operation amplifies higher values and suppresses lower values. This step might be used to enhance certain features of interest within the signal, or it could be related to calculating the energy or power of the signal. Figure 2d—Moving Window Integrated Signal Output: This Fig likely involves applying a moving window integration to the signal. Moving window integration involves summing or averaging signal values within a sliding window. This can help in smoothing out the signal and extracting underlying trends or patterns, while also

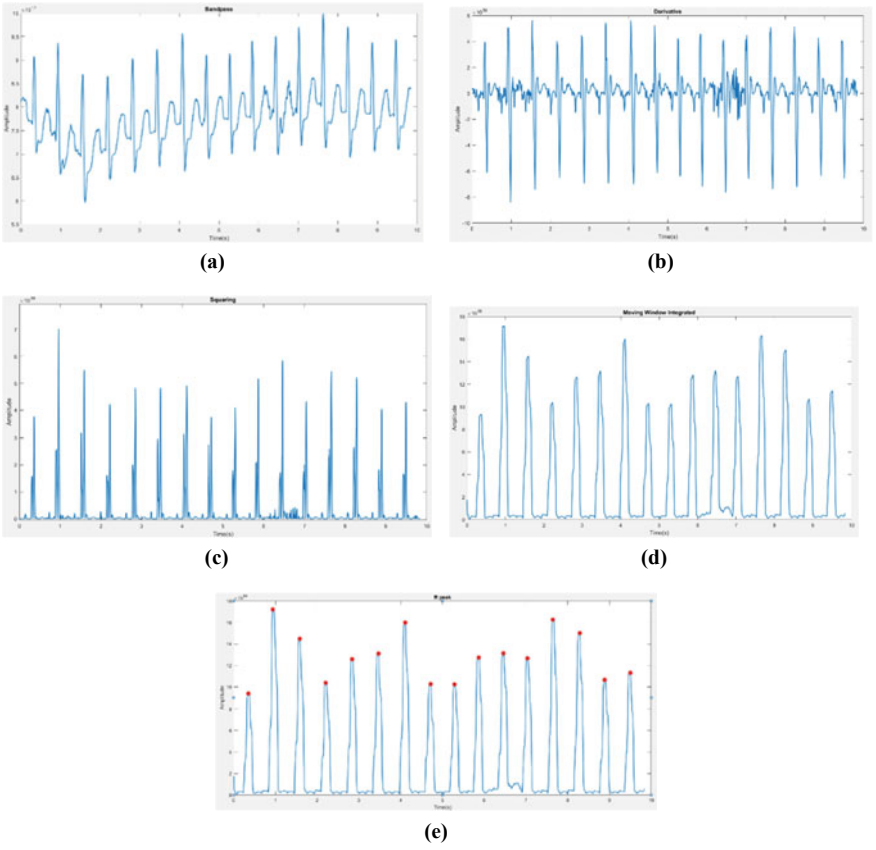
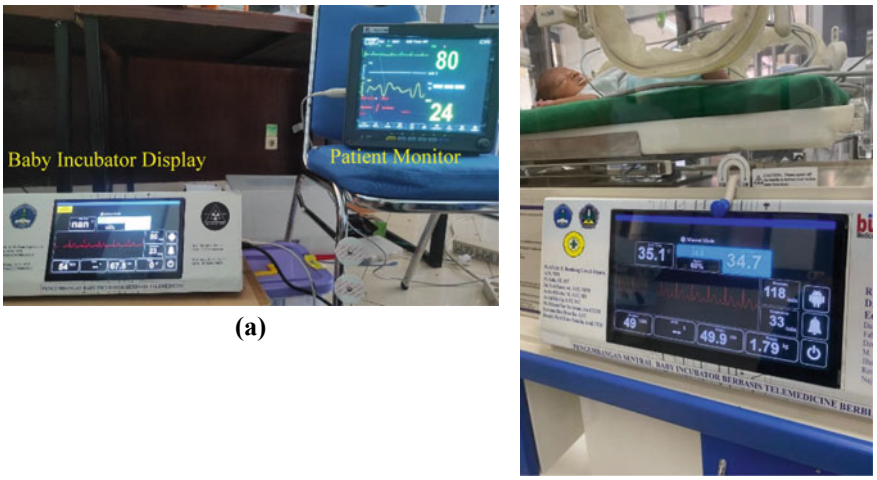


Fig. 2 Is the result of testing the pantomkin method with **a** which is a band pass filter process. **b** Derivative signal output. **c** Squaring signal output. **d** Moving Window Integrated signal output. **e** and R Peak detection

reducing noise. Figure 2e —R Peak Detection: The final step appears to involve detecting R-peaks. In the context of physiological data, R-peaks correspond to the highest points of the QRS complex in an electrocardiogram (ECG) signal, representing the depolarization of the ventricles in the heart. Detecting R-peaks is a crucial step in heart rate analysis and arrhythmia detection. Overall, the process seems to be a sequence of signal processing techniques aimed at enhancing the relevant features in the input data while reducing noise. This kind of processing is often seen in fields like biomedical signal analysis, where accurate feature extraction and noise reduction are essential for subsequent analysis and interpretation. In Table 1 it is explained about the results of measurements that have been carried out on subjects with an average value of 5 times data collection.

From testing and measurements, the results obtained were different variations in error differences for each respondent, namely between 0.2 and 0.4 BPM for HR



(a) (b)

Fig. 3 Baby incubator test results **a** comparison of the display on the baby incubator display with the patient monitor display. **b** monitoring the baby's condition using a baby incubator equipped with an ECG signal display and the number of heart rate

Table 1 HR measurement result in subject

Subject	Module average score (Bpm)	Patient monitor average value (Bpm)	Error ±
1	118,6	118,4	0,2
2	120,4	120,8	−0,4
3	115,8	115,4	0,4
4	118,4	118	0,4
5	120,4	120,8	−0,4
6	115,8	115,4	−0,4
7	119,2	119,6	−0,4
8	120,4	120,8	−0,4
9	116,8	116,4	0,4
10	118,4	118,8	−0,4

parameters with the largest difference obtained being 0.4. Figure 3 is an explanation of the results of measuring a baby's heart rate using a baby incubator which is equipped with ECG signal monitoring as well as the number of heart beats.

In Fig. 3 illustrates the results obtained from measurements using a module that shows the difference in results that are not far from the allowable tolerance limits so that it meets the standard. The comparison results between the baby incubator display and the patient monitor are explained in Fig. 4.

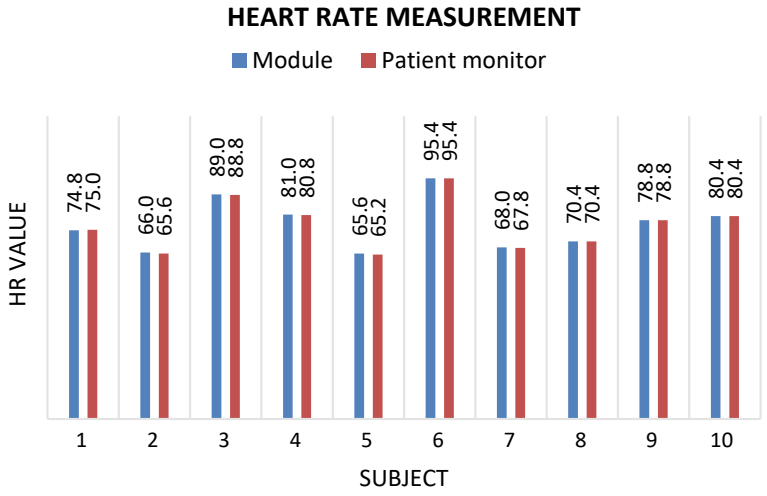


Fig. 4 HR value from measurement

Table 2 T TEST: equal variances

T TEST: equal variances				Alpha	0,05			
	Std err	t-stat	df	p-value	t-crit	Lower	Sig	Effect r
One tail	4,494,881	0,040,045	18	0,48,424	1,73,406		No	0,0094
Two tail	4,494,881	0,040,045	18	0,96,849	2,10,092	−9,62,339	No	0,0094

Based on the test results in Table 2, we do not have enough evidence to reject the null hypothesis. A p-value greater than the alpha level of significance (0.05) indicates that there is no significant difference between the two groups tested. A T-Statistic lower than the Critical T value also indicates the statistical insignificance of the difference. A low effect size (effect r) indicates that the difference between the groups is also small.

From Table 2 it is explained that the results of statistical calculations obtained a p value = 0,48,424 which means that $p > 0.05$.

4 Discussion

The results obtained from testing and measurement use the Pan-Tompkins method, where variations in error differences are observed for different respondents. Specifically, errors in heart rate (HR) measurements ranged from 0.2 to 0.4 beats per minute (BPM), with the largest difference being 0.4 BPM. This variation represents the difference between the HR value calculated using the Pan-Tompkins method and

the actual HR value. A difference of 0.2–0.4 BPM indicates that there may be slight inaccuracies in the Pan-Tompkins method measurements compared to actual HR values. These differences can be caused by a variety of factors, including noise in the ECG signal, variations in individual heart rhythms, or limitations in the algorithm itself. It is important to note that a difference of 0.2–0.4 BPM may be considered acceptable, depending on the specific application and level of accuracy required. For example, in many clinical applications, such small differences may not have a significant impact on the interpretation of the results. However, for highly sensitive applications or research studies that require very precise HR measurements, further method refinements or additional signal processing techniques may be necessary to minimize these differences. A p-value greater than 0.05 in a statistical test usually indicates that there is no significant difference between the compared groups. The result obtained is a p value greater than 0.05, meaning there is no significant difference between the HR value calculated using the Pan-Tompkins method and the actual HR value. When $p > 0.05$ indicates weak evidence against the null hypothesis, which means you failed to reject the null hypothesis.

$p \leq 0.05$ indicates strong evidence against the null hypothesis, which means you would reject the null hypothesis. In your case, $p > 0.05$, you do not have enough evidence to conclude that the difference in HR values between the Pan-Tompkins method and the true values is statistically significant. These results indicate that any observed variations or differences between calculated HR values and actual values are most likely due to random chance or variability inherent in the data, and not a result of the method itself.

In other studies that monitor vital signs in infant incubators at heart rate values using non-contact methods get results that are 3% but the results are not displayed directly which must be processed first [27]. The result of this study is the monitoring of baby incubators with vital sign ECG using the pan tompkins method in real-time successfully used on heart rate measurement parameters and obtaining an overall error value of ± 0.4. Although getting good results there are some shortcomings including the results of the graph displayed there is still noise and the performance of the program feels quite heavy, so in our opinion it can still be optimized to get better results. The results of this study can be used for doctors to monitor the condition of premature babies remotely using a smart phone or computer so that it is expected to accelerate the analysis and treatment of patients.

5 Conclusions

This research aims to develop a telemedicine-based baby incubator with accurate ECG monitoring that can facilitate monitoring of the baby's condition directly so as to get the right treatment. The program on the arduino can manage all work systems on the device and the results will be forwarded to the raspberry for ECG signal processing and IOT settings. The output of the ECG circuit in lead II which has been processed in the arduino microcontroller will be sent to the raspberry and ECG signal

processing is carried out using the Pan-Tompkins algorithm in the phyton program so that the QRS complex is obtained as well as the peak of the R signal which is then calculated into a HR value. After testing the module measurements with the patient monitor comparison, there is the smallest difference value of 0, the largest difference value is 0.4 and has an overall mean of ±0,4, so it can be concluded that the measurement of our tool is suitable for use because it is still below the allowable error value. The results of the graphs displayed still have noise that is expected to come from the patient's body as well as the source voltage mesh used. In further research it is recommended to try using other calculation methods or formulas to reduce the noise that still exists besides that, it can also use mini computers (raspberries, etc.) with better specifications, and there is a need for optimization of the phyton program used so that it is not too heavy in its performance.

References

1. Kristya F, Luthfiyah S, Hari Wisana IDG, Thaseen M (Jan 2021) Baby incubator monitoring center for temperature and humidity using WiFi network. J Electron, Electromed Eng Med Inform 3(1):8–13. https://doi.org/10.35882/jeeemi.v3i1.2
2. Shaib M, Rashid M, Hamawy L, Arnout M, El Majzoub I, Zaylaa AJ (Oct 2017) Advanced portable preterm baby incubator. In: 2017 fourth international conference on advances in biomedical engineering (ICABME). IEEE. pp 1–4. https://doi.org/10.1109/ICABME.2017.8167522
3. Lamidi L, Kholiq A, Ali M (2021) A low cost baby incubator design equipped with vital sign parameters. Indones J Electron Electromed Eng Med Inform 3(2):53–58. https://doi.org/10.35882/ijeeemi.v3i2.3
4. Silveira MF et al (Apr 2019) Low birthweight and preterm birth: trends and inequalities in four population-based birth cohorts in Pelotas, Brazil, 1982–2015. Int J Epidemiol 48(Supplement_1):i46–i53. https://doi.org/10.1093/ije/dyy106
5. Greer R, Olivier C, Pugh JE, Eklund JM, McGregor C (Aug 2014) Remote, real-time monitoring and analysis of vital signs of neonatal graduate infants. In: 2014 36th annual international conference of the IEEE engineering in medicine and biology society. IEEE, pp 1382–1385. https://doi.org/10.1109/EMBC.2014.6943857
6. Chen W, Bambang Oetomo S, Feijs L, Bouwstra S, Ayoola I, Dols S (Oct 2010) Design of an integrated sensor platform for vital sign monitoring of newborn infants at neonatal intensive care units. J Healthc Eng 1(4):535–553. https://doi.org/10.1260/2040-2295.1.4.535
7. Wahab MA, Md Nor D (2021) Safety and health monitoring system for baby incubator using IoT. Evol Electr Electron Eng 2(2):256–264. https://doi.org/10.30880/eeee.2021.02.02.031
8. Aktaş F, Aktas F, Kavus E, Kavus Y (2017) A real-time infant health monitoring system for hard of hearing parents by using android-based mobil devices IoT-based healthcare framework for biomedical applications view project a real-time infant health monitoring system for hard of hearing parents by using android-based mobil devices. https://www.researchgate.net/publication/315696973
9. Young A, Azeez F, Godad SP, Shetty P, Sharma A (2021) A multimodal quality improvement approach to promote normothermia in very preterm infants. Acta Paediatr 110(10):2745–2752. https://doi.org/10.1111/apa.16009
10. Carns J et al (2018) Impact of hypothermia on implementation of CPAP for neonatal respiratory distress syndrome in a low-resource setting. PLoS ONE 13(3):e0194144. https://doi.org/10.1371/journal.pone.0194144

11. Zamora GAA, Garcia DC (Dec 2021) A prospective study on the efficacy and safety of kangaroo mother care as an alternative means of transport of preterm and term small-for-gestational age infants. Acta Med Philipp 55(9). https://doi.org/10.47895/amp.v55i9.3898
12. Jagadish AS, Benakappa A, Benakappa N, Morgan G (Dec 2019) A randomized control trial of hypothermia alert device in low birth weight newborns and the effect on kangaroo mother care and weight gain. Int J Contemp Pediatrics 7(1):52. https://doi.org/10.18203/2349-3291.ijcp20195725
13. Rifat Hazari M, Jahan E, Rashedul Amin M, Alam MA (May 2014) Design and implementation of heart rate measurement device using wireless system. In: 2014 international conference on informatics, electronics & vision (ICIEV). IEEE, pp 1–5. https://doi.org/10.1109/ICIEV.2014.6850732
14. Das N, Chakraborty M (Nov 2017) Performance analysis of FIR and IIR filters for ECG signal denoising based on SNR. In: 2017 third international conference on research in computational intelligence and communication networks (ICRCICN). IEEE, pp 90–97. https://doi.org/10.1109/ICRCICN.2017.8234487
15. Linde JE et al (2016) Normal newborn heart rate in the first five minutes of life assessed by dry-electrode electrocardiography. Neonatology 110(3):231–237. https://doi.org/10.1159/000445930
16. Chen W, Nguyen ST, Coops R, Oetomo SB, Feijs L (Nov 2009) Wireless transmission design for health monitoring at neonatal intensive care units. In: 2009 2nd international symposium on applied sciences in biomedical and communication technologies. IEEE, pp 1–6. https://doi.org/10.1109/ISABEL.2009.5373671
17. Subramanian M, Sheela T, Srividya K, Arulselvam D (Aug 2019) Security and health monitoring system of the baby in incubator. Int J Eng Adv Technol 8(6):3582–3585. https://doi.org/10.35940/ijeat.F9353.088619
18. Manlong NA, Rahul J, Sora M (2018) ST segment analysis for early detection of myocardial infarction. Int J Comput Sci Eng 6(6):1500–1504. https://doi.org/10.26438/ijcse/v6i6.15001504
19. Qin Q, Li J, Yue Y, Liu C (2017) An adaptive and time-efficient ECG R-peak detection algorithm. J Healthc Eng 2017. https://doi.org/10.1155/2017/5980541
20. Jenkal W, Latif R, Toumanari A, El B'Charri O (2016) Efficient method of QRS complex extraction using a multilevel algorithm and an adaptive thresholding technique. IKn: Proceedings of 2015 IEEE world conference on complex systems, WCCS 2015. https://doi.org/10.1109/ICoCS.2015.7483281
21. Vastarouchas C, Kapoulea S, Psychalinos C (Dec 2016) ECG signal acquisition for the Pan-Tompkins algorithm using current-mirror filters. In: 2016 IEEE international conference on electronics, circuits and systems (ICECS). IEEE, pp 317–320. https://doi.org/10.1109/ICECS.2016.7841196
22. Lu X, Pan M, Yu Y (2018) QRS detection based on improved adaptive threshold. J Healthc Eng 2018. https://doi.org/10.1155/2018/5694595
23. Liu F et al (2017) The accuracy on the common Pan-Tompkins based QRS detection methods through low-quality electrocardiogram database. J Med Imaging Health Inform 7(5):1039–1043. https://doi.org/10.1166/jmihi.2017.2134
24. Wu X, Wang Z, Xu B, Ma X (2020) Optimized Pan-Tompkins based heartbeat detection algorithms. In: Proceedings of the 32nd Chinese control and decision conference, CCDC 2020, pp 892–897. https://doi.org/10.1109/CCDC49329.2020.9164736
25. Rana A, Kim K (2019) Cardiac disease detection using modified Pan-Tompkins algorithm. J Sens Sci Technol 28(1):13–16
26. Irianto BG, Maghfiroh AM, Sofie M, Kholiq A (2023) Baby incubator with overshoot reduction system using PID control equipped with heart rate monitoring based on internet of things. Int J Technol 14(4):811. https://doi.org/10.14716/ijtech.v14i4.5678

27. Klaessens JH, van den Born M, van der Veen A, Sikkens-van de Kraats J, van den Dungen FA, Verdaasdonk RM (Feb 2014) Development of a baby friendly non-contact method for measuring vital signs: first results of clinical measurements in an open incubator at a neonatal intensive care unit, Vo-Dinh T, Mahadevan-Jansen A, Grundfest WS (eds), p 89351P. https://doi.org/10.1117/12.2038353

Recent Advances and Challenges in 3D Printing of Prosthetic Hands



Triwiyanto, Sari Luthfiyah, Bedjo Utomo, I. Putu Alit Pawana, Wahyu Caesarendra, and Vijay Anant Athavale

Abstract Traditional methods of creating prosthetics are often expensive, time-consuming, and lack personalization. However, 3D printing technology has emerged as a transformative force that can address these challenges. The purpose of this abstract is to analyze the research documents on this topic from Scopus, a database of peer-reviewed literature and scientific information. The search query was TITLE-ABS-KEY (prosthetic AND hand AND 3d AND printing) AND PUBYEAR > 2012 AND PUBYEAR < 2024, which returned 236 documents in total. The document analysis reveals that the research on prosthetic hands using 3D printing technology has increased over time, especially in 2020 and 2022. In conclusion, the research on prosthetic hands using 3D printing technology is a growing and diverse field that involves many researchers, institutions, countries/territories, and disciplines. The use of various filaments such as PLA, TPU, and silicone rubber has enabled the creation of low-cost, functional, and customizable prosthetic devices. However, there is still room for improvement and further research in this field to continue advancing this important technology.

Keywords Prosthetic hand · 3D printing · Recent advances

Triwiyanto (✉) · S. Luthfiyah · B. Utomo
Department of Medical Electronics Technology, Poltekkes Kemenkes Surabaya, Surabaya, Indonesia
e-mail: triwiyanto@ieee.com

I. P. A. Pawana
Dr. Soetomo Hospital, Rehabilitation Unit, Surabaya, Indonesia

W. Caesarendra
Universiti Brunei Darussalam, Bandar Seri Begawan, , Brunei Darussalam

V. A. Athavale
Walchand Institute of Technology, Solapur, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

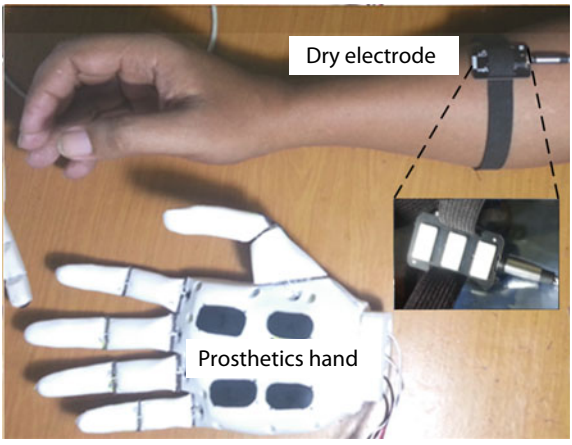
T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_43

1 Introduction

The advent of 3D printing technology has revolutionized various fields, including the development of prosthetics [12, 27, 35, 47]. This essay aims to provide a comprehensive overview of the research on prosthetic hands using 3D printing technology. Prosthetic hands are artificial devices that replace or enhance the function of missing or impaired hands. They can improve the quality of life, mobility, and independence of people with limb loss or impairment [13, 43]. However, traditional methods of creating prosthetic hands are often expensive, time-consuming, and lack personalization. 3D printing technology offers a low-cost, fast, and customizable way to create prosthetic hands that can meet the specific needs and preferences of individual users [20, 24, 29]. 3D printing technology is a process that creates physical objects from digital models by depositing layers of material on top of each other. This technology has many advantages for prosthetic hand design. For instance, it can create complex and intricate shapes that are difficult or impossible to achieve with conventional methods [40, 51]. It can use various materials, such as plastics, metals, ceramics, or composites, to create prosthetic hands with different properties and functionalities [4, 46]. It can reduce the material waste and environmental impact compared to traditional manufacturing methods [17, 18, 23]. Furthermore, it can enable mass customization and personalization of prosthetic hands by using digital scanning and modeling techniques [3, 8].

The state-of-the-art method of prosthetic hand design using 3D printing technology involves several steps. The first step is capturing the user's hand geometry and motion data using sensors, cameras, or scanners [3, 34]. The next step is designing the prosthetic hand model using computer-aided design (CAD) software or generative design algorithms [28]. After that, the appropriate material and 3D printing technique are selected based on the desired characteristics and performance of the prosthetic hand [44]. The prosthetic hand parts are then printed layer by layer using a 3D printer [2]. Finally, the prosthetic hand device is assembled and tested, and adjustments are made to fit the user's needs and preferences [6, 22, 45]. In conclusion, prosthetic hand design using 3D printing technology is a promising field that has many applications in various domains. The use of various filaments such as PLA, TPU, and silicone rubber has enabled the creation of low-cost, functional, and customizable prosthetic devices (Fig. 1). The state-of-the-art method of prosthetic hand design involves capturing, designing, selecting, printing, assembling, and testing steps that can produce high-quality and personalized devices. The purpose of this review paper is to provide a comprehensive overview of the research on prosthetic hand design using 3D printing technology. The contribution of this review paper is to:

1 Provide a comprehensive overview of the research on prosthetic hand design using 3D printing technology based on a search query from Scopus, a database of peer-reviewed literature and scientific information.
2 Summarize and analyze the search results by different criteria, such as year, source, author, affiliation, country/territory, type, subject area, and funding sponsor, using various charts and tables.



(a)



(b)

Fig. 1 Prosthetic hand design using 3D printing model **a** flexible filaments [50], and **b** PLA filaments [49]

3 Highlight the advantages and challenges of 3D printing technology for prosthetic hand design, such as low-cost, functionality, customization, complexity, durability, and personalization.

2 Method

The search query was TITLE-ABS-KEY (prosthetic AND hand AND 3d AND printing) AND PUBYEAR > 2012 AND PUBYEAR < 2024, which means that the search was looking for documents that have the words prosthetic, hand, 3d, and printing in their title, abstract, or keywords, and that were published between 2013 and 2023. The search returned 236 documents in total. The result shows various charts and tables that summarize and analyze the search results (Fig. 2). The first chart shows the number of documents per year, which indicates that the research on prosthetic hands using 3D printing technology has increased over time, especially in 2020 and 2022. The fourth chart shows the number of documents per country/territory, which indicates that the United States, India, Australia, United Kingdom,

and China are the top five countries/territories that produce the most documents on this topic. The sixth chart shows the number of documents per type, which shows that most of the documents are articles or conference papers, followed by reviews and book chapters. The seventh chart shows the number of documents per subject area, which demonstrates that engineering, computer science, medicine, materials science, and physics and astronomy are the main subject areas that cover this topic. The eighth chart shows the number of documents per funding sponsor, which reveals that Engineering and Physical Sciences Research Council, European Regional Development Fund, National Science Foundation, National Natural Science Foundation of China, and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior are some of the major funding agencies that support this research. the charts provide a comprehensive overview of the research on prosthetic hands using 3D printing technology based on a search query from Scopus.

The recent analysis of 236 documents from the Scopus database, using the search strategy "prosthetic AND hand AND 3D and printing" and visualized using the VOSVIEWER application, has revealed four distinct clusters of keywords (Fig. 3). These clusters provide a fascinating insight into the current trends and focus areas in the field of 3D printed prosthetic hands. The first cluster is centered around the technology and design aspects of 3D printed prosthetic hands. Keywords such as "3-d printing", "3d modeling", "computer aided design", and "design" highlight the importance of technology in creating these prosthetics. The presence of terms like "anthropomorphic hand", "artificial limbs", "biomimetics", and "bionics" indicates a focus on creating prosthetics that closely mimic the form and function of natural limbs. Interestingly, this cluster also includes keywords like "cost effectiveness" and "developing countries", suggesting a focus on making these advanced prosthetics more accessible and affordable. The second cluster appears to be more focused on the practical application and usability of these prosthetics. Keywords such as "activities of daily living", "range of motion", and "upper limb" suggest a focus on ensuring that these prosthetics are functional and useful in everyday life. The presence of terms like "adult", "female", and "case report" indicates that these studies are not just theoretical but are based on real-world cases and experiences [25]. The third cluster is centered around the control systems used in these prosthetics. Keywords like "biomedical signal processing", "control systems", "electromyography", and "myoelectric control" suggest a focus on developing advanced control systems that allow users to control their prosthetics in a natural and intuitive way. The presence of terms like "child" and "young adult" indicates that these control systems are being designed with users of all ages in mind. The fourth cluster seems to be focused on the biomechanics of these prosthetics. Keywords such as "arm", "finger", "hand strength", and "thumb" suggest a focus on understanding and replicating the biomechanics of the human hand. The presence of terms like "machine learning" and "soft robotics" indicates an interest in using advanced technologies to improve the functionality and adaptability of these prosthetics.

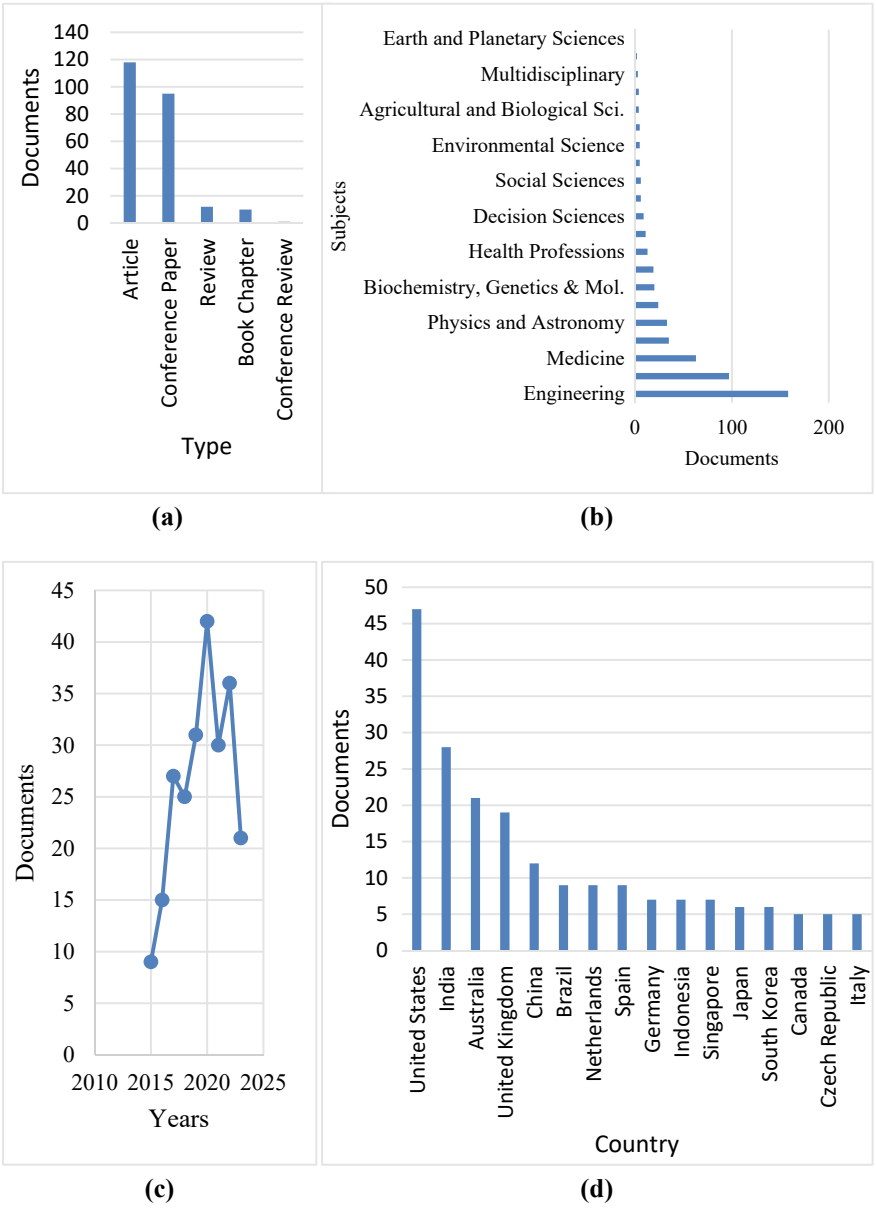


Fig. 2 The analyses of **a** type of article, **b** subject, **c** years, and **d** country from Scopus.com database

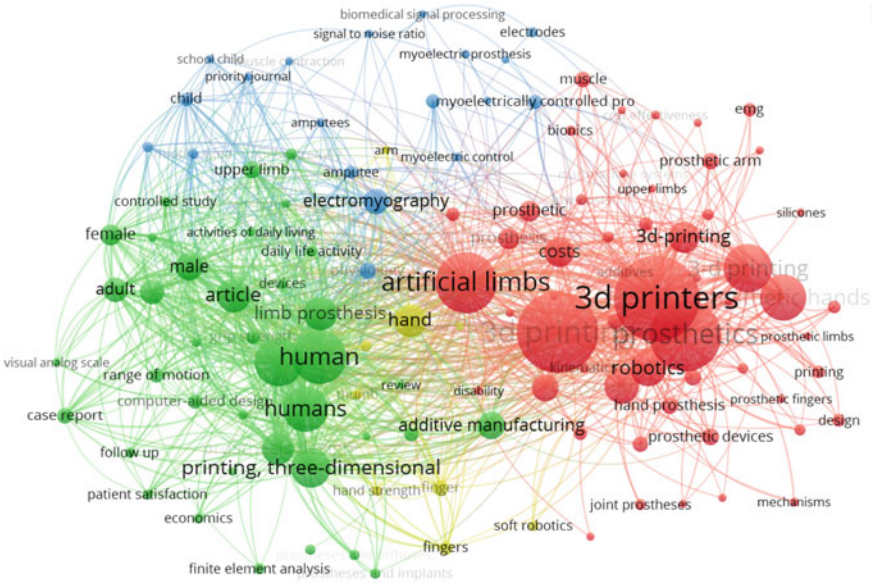


Fig. 3 Clusters map based on bibliography data using Scopus.com database

3 Result and Discussion

3D printing is a rapidly evolving technology that has various applications in different fields such as engineering, medicine, dentistry, education, etc. This paper reviews some of the recent research articles that use 3D printing techniques to create innovative products and solutions for various purposes and challenges [42]. The paper summarizes the main methods, findings, performance, and limitations of each article and provides a comparative analysis of their contributions and implications. The paper aims to provide a comprehensive overview of the current state-of-the-art and future trends of 3D printing technologies and their impact on society (Table 1).

3D printing, also known as additive manufacturing, is a process that creates three-dimensional objects by depositing materials layer by layer according to a digital model. 3D printing has various applications in biomedical engineering, such as producing implants, prosthetics, devices, structures, and artificial organs and tissues. This essay reviews some of the recent research articles that use 3D printing techniques to create innovative products and solutions for various biomedical challenges. One of the applications of 3D printing is to create orthotics and prosthetics for animals with disabilities or injuries [26]. Mendaza-DeCal et al. developed a low-cost and fast method to produce customized devices for veterinary applications using a novel fabrication flow [14, 37]. They used 3D scanning, modeling, and printing techniques to create orthoses and prostheses for dogs, cats, horses, and birds. They reported positive feedback from veterinarians and animal owners on the functionality and usability of the devices (Table 2).

Table 1 Comprehensive analysis of keyword clusters in 3D printed prosthetic hand research: unveiling current trends and future directions

Cluster	Main keyword
1	3-d printing, 3d modeling, 3d printers, anthropomorphic hand, artificial limbs, biomimetics, bioninc arm, bionics, computer aided design, controlles, cost effectiveness, design, developing countries, emg, hand prosthesis, joint prostheses, low-cost
2	Activites of daily living, adult, arm amputation, case report, article, clinical article, devices, equipment design, female, printing 3-dimension, prosthesis design, range of motion, upper limb, 3D printing
3	Amputee, biomedical signal processing, child, control systems, electrodes, electromyography, muscle contraction, myoelectric control, myoelectric prosthesis, Myo electrically controlled, surface electromyography, wearable sensors, young adult
4	Arm, biomechanics, finger, hand, hand strength, machine learning, soft robotics, thumb

Another application of 3D printing is to create ear prostheses for patients with ear defects or deformities. Cruz et al. proposed a new approach to generate realistic ear models based on geometric parameters and 3D scans of the intact ear [15]. They used an automated parametric ear model and frugal 3D scanning methods to create high-quality prosthetic ears using 3D printing techniques. They demonstrated the feasibility and accuracy of their method using a phantom head and a volunteer subject. A third application of 3D printing is to assess the knowledge, awareness, and practices of dental professionals regarding the use of 3D printing in dentistry. Acharya et al. conducted two cross-sectional surveys among dental practitioners and dental laboratory technicians in Karnataka, India [1]. They found that most of the respondents had low to moderate knowledge, awareness, and practices regarding 3D printing in dentistry. They identified the barriers and challenges faced by the dental professionals in adopting 3D printing technologies, such as lack of training, equipment, materials, quality control, regulation, etc. A fourth application of 3D printing is to design and fabricate low-cost tendon prosthetic hands for developing countries. Andrés-Esperanza et al. designed and fabricated prosthetic hands that met the expectations of potential users in terms of functionality, weight, cost, and appearance [3, 11]. They used 3D printing techniques to create the hand structure and components, such as fingers, palm, wrist, tendons, etc. They compared their devices with existing commercial products in terms of functionality, weight, cost, and appearance. A fifth application of 3D printing is to develop a novel methodology for multi-material design and fabrication of medical assistive devices. Faller et al. developed a new workflow to create complex geometries and material combinations for medical assistive devices using 3D printing techniques [5, 10]. They presented some examples of their devices, such as orthoses, prostheses, implants, and surgical instruments. They used different types of materials, such as polymers, metals, ceramics, composites, etc., to achieve different functions and properties. A sixth application of 3D printing is to create artificial organs and tissues using

Table 2 A summary of the methods, findings, performance, and limitations of prosthetic hand design using 3D models

No	Author	Method	Finding	Limitation
1	Andrés-Esperanza et al. 2022 [3]	Experimental evaluation of 3D-printed tendon prosthetic hands	The hands are affordable, functional and customizable, but have low user satisfaction and acceptance	The study only tested one hand model and did not compare it with other prosthetic options
2	Yan et al. 2022 [51]	Design, kinematic modeling and evaluation of a novel soft prosthetic hand with abduction joints	The hand can achieve adaptive grasping and dexterous manipulation with low actuation power and high flexibility	The hand has limited payload capacity and durability, and needs further optimization of the design parameters and control strategy
3	Dhir et al. 2023 [9]	Comparative analysis of 3D printing in large, medium and small firms using survey data and interviews	The firms have different motivations, challenges and outcomes of adopting 3D printing, depending on their size and industry sector	The study has a small sample size and does not account for the dynamic nature of 3D printing adoption over time
5	Hummelink et al. 2021 [22]	Designing and utilizing 3D-printed skin incision guides during the first Dutch bilateral hand-arm transplantation	The guides improved the accuracy, efficiency and safety of the surgical procedure and reduced the risk of infection and bleeding	The study only reports one case and does not compare the outcomes with conventional methods or other types of guides
7	Setty et al. 2020 [46]	Design Considerations of the Touch Hand 4 using CAD software and finite element analysis	The hand is lightweight, robust and modular, but has high power consumption and limited grip force	The study only presents the design process and does not evaluate the performance or functionality of the hand in real-world scenarios

(continued)

Table 2 (continued)

No	Author	Method	Finding	Limitation
8	Llop-Harillo et al. [34]	The Anthropomorphic Hand Assessment Protocol (AHAP)	The protocol is a standardized and comprehensive tool to evaluate the functionality and performance of anthropomorphic robotic hands	The protocol does not consider the user's preferences, satisfaction or comfort, and may not be applicable to non-anthropomorphic hands
9	León-Camacho et al. [31]	Innovation of robotic hand with two thumbs	The hand can perform more stable and versatile grasps than conventional hands, but has higher complexity and cost	The hand has not been tested with real users or in real-world scenarios, and may have limited compatibility with existing prosthetic sockets or interfaces
10	Hajra et al. [16]	Revolutionizing self-powered robotic systems with triboelectric nanogenerators (TENGs)	TENGs can harvest ambient mechanical energy to power robotic systems without batteries or wires, but have low output voltage and current compared to other energy harvesting devices	TENGs have not been widely integrated with robotic systems due to challenges such as scalability, durability and compatibility with existing electronics or sensors
11	Hernandez-Hinojosa et al. [21]	Bioinspired robotic exoskeleton for endotracheal intubation (ETI)	The exoskeleton can assist medical personnel in performing ETI with high precision and safety, but has high design complexity and fabrication cost compared to conventional ETI devices	The exoskeleton has not been tested with human subjects or in clinical settings, and may have ethical or legal implications for its use in medical procedures

(continued)

Table 2 (continued)

No	Author	Method	Finding	Limitation
12	Herath et al. [19]	The interface can control the robot fingers using electroencephalography (EEG) signals from the user's brain, but has low accuracy and reliability compared to other control methods	The interface only works with one type of robot finger (bio-inspired) and one type of EEG signal (motor imagery), and may have issues such as noise, latency or user fatigue	Brain-computer interfacing for flexion and extension of bio-inspired robot fingers
13	Li et al. [33]	Fabrication and testing of flexible and wearable healthcare sensors for visual reality health-monitoring	The sensors can monitor the vital signs and physiological parameters of the user in real time and provide feedback through visual reality devices, but have challenges such as power consumption, data security and user comfort	The sensors only work with specific types of visual reality devices and have not been evaluated in clinical settings or with large-scale users
14	Leggieri et al. [30]	Design and testing of a dexterous gripper with embedded object localization algorithm	The gripper can perform in-hand manipulation of objects with different shapes and sizes, but has high computational complexity and sensitivity to noise	The gripper only works with specific types of objects and has not been evaluated in real-world scenarios or with human–robot interaction

bioprinting techniques. Bioprinting is a form of 3D printing that uses cells, tissues, growth factors, biomaterials and other biological related materials as building blocks to fabricate living structures [36]. Bioprinting has been used to create various types of organs and tissues, such as skin, bone, cartilage, blood vessels, liver, kidney, heart, etc., for transplantation, drug testing, disease modeling, and tissue engineering purposes [39, 52]. Bioprinting has several advantages over conventional tissue engineering methods, such as high precision, customization, scalability, vascularization, and integration. These are some of the examples of how 3D printing technologies can

Table 3 3D printing filament to design a prosthetic hand

No	Author	Filament type	Object	Performance
1	Andrés-Esperanza et al. [3]	PLA and TPU	Tendon prosthetic hands	Low cost, high functionality, and user satisfaction
2	Yan et al. [51]	TPU and silicone rubber	Soft prosthetic hand with abduction joints	High flexibility, dexterity, and grasping ability
4	Siraj and Bharti [47]	PLA and ABS	Various parts and models	Reliability assessment based on failure modes and effects analysis
5	Hummelink et al. [22]	PLA and PVA	Skin incision guides for hand-arm transplantation surgery	High accuracy, precision, and safety
6	Saeed et al. [44]	PLA with continuous carbon fiber reinforcement	Composite plates with different fiber orientations	Prediction of mechanical properties using classical laminated-plate theory
7	Setty et al. [46]	PLA and ABS with carbon fiber reinforcement	Touch Hand 4 prosthetic hand prototype	Improved design, functionality, and durability
8	Haleem and Javaid [17]	PLA, ABS, nylon, PEEK, etc	Medical parts such as implants, surgical guides, etc	Advantages and challenges of additive manufacturing in healthcare
9	Tom et al. [48]	Biocompatible materials such as PCL, PLA, etc	Biomedical devices such as scaffolds, implants, etc	Recent research developments and applications of additive manufacturing in biomedicine
10	Ali et al. [2]	PLA and ABS with different colors and diameters	Large-scale parts and models with multi-colors and multi-materials	Development and evaluation of a novel multi-extrusion FDM printer
11	Ntagios et al. [41]	Various inks such as graphene, silver, etc	Direct ink extruder system with multi-part materials mixing	Closed-loop control, high resolution, and versatility
13	León-Camacho et al. [31]	PLA and ABS	Robotic hand with two thumbs	Increased dexterity, stability, and manipulation capabilities
14	Muminović et al. [38]	PLA	Polylactic acid gears with different geometries and loads	Experimental failure analysis using digital image correlation

(continued)

Table 3 (continued)

No	Author	Filament type	Object	Performance
15	Leung et al. [32]	PLA and PVA	Multi-articulated joints of non-assembly mechanisms such as grippers and linkages	Optimization of support structure design and fabrication parameters
17	Kuipers et al. [28]	PLA with different colors and diameters	Interlaced topologically interlocking lattice structures with dual-material extrusion	Improved mechanical properties, aesthetics, and functionality
19	Hernandez-Hinojosa et al. [21]	PCL and silicone rubber with embedded sensors and actuators	Bioinspired robotic exoskeleton for endotracheal intubation	High flexibility, accuracy, and safety

be used for biomedical applications [7]. The advantages of 3D printing technologies include high precision, flexibility, customization, personalization, innovation, etc. However, there are also some challenges and limitations that need to be addressed before 3D printing technologies can be widely adopted for biomedical applications. Some of these challenges include biocompatibility, sterilization, regulation, quality control, cost-effectiveness, patient acceptance, etc. The 3D filament materials are very important to design a prosthetic hand. Table 3 shows the type of filament used and the performance.

3D printing technology is a rapidly evolving field that has many applications in various domains, such as engineering, medicine, biotechnology, robotics, etc. The papers in the tables present some of the recent research developments and challenges in 3D printing technology and its impact on society. The papers can be grouped into four main categories: prosthetics, medical parts, composite materials, and large-scale models. The first category is prosthetics, which are artificial devices that replace or enhance the function of missing or impaired body parts. 3D printing technology offers a low-cost, customizable, and accessible way to create prosthetic devices that can improve the quality of life of people with disabilities. Some of the papers in this category focus on the design and evaluation of prosthetic hands that can mimic the natural movements and sensations of human hands. For example, Javier Andrés-Esperanza et al. developed tendon prosthetic hands using PLA and TPU filaments, which were praised for their low cost, high functionality, and user satisfaction [3, 43]. Similarly, Yadong Yan et al. used TPU and silicone rubber to create a soft prosthetic hand with abduction joints, which demonstrated high flexibility, dexterity, and grasping ability [51]. Kiran Setty et al. improved the design, functionality, and durability of the Touch Hand 4 prosthetic hand prototype using PLA and ABS with carbon fiber reinforcement [46]. In the field of medical parts production, S. Hummelink et al. used PLA and PVA to create skin incision guides for hand-arm transplantation surgery, which were noted for their high accuracy, precision, and safety [14, 22]. Abid Haleem and

Mohd Javaid discussed the advantages and challenges of using various filaments like PLA, ABS, nylon, PEEK, etc., in healthcare for creating implants, surgical guides, etc. [23]. Large-scale parts and models with multi-colors and multi-materials were developed by Ali et al. using PLA and ABS with different colors and diameters [2]. Tim Kuipers et al. created interlaced topologically interlocking lattice structures with dual-material extrusion that improved mechanical properties, aesthetics, and functionality [28]. In terms of composite materials fabrication, Saeed et al. used PLA with continuous carbon fiber reinforcement to create composite plates with different fiber orientations. They predicted the mechanical properties using classical laminated-plate theory [44]. In conclusion, 3D printing technology is a promising field that has many applications in various domains. The papers in the tables present some of the recent research developments and challenges in 3D printing technology and its impact on society. The papers differ in their filament types, objects, and performance measures, but they share a common theme of 3D printing technology and its innovation (Table 3).

4 Future Direction

3D printing technology has shown great potential in the development of prosthetic hands. The use of various filaments, such as PLA, TPU, and silicone rubber, has enabled the creation of low-cost, functional, and customizable prosthetic devices. However, there is still room for improvement and further research in this field. One future direction for the development of prosthetic hands using 3D printing models is to improve the functionality and dexterity of the devices. This can be achieved by incorporating more advanced sensors and actuators into the design, allowing for more natural movements and sensations. Additionally, the use of more flexible and durable materials, such as carbon fiber reinforcement or biocompatible materials, can improve the durability and comfort of the prosthetic hands. Another area for future research is to improve the customization and personalization of prosthetic hands. 3D printing technology allows for the creation of prosthetic devices that are tailored to the specific needs and anatomy of individual users. By incorporating more advanced scanning and modeling techniques, it may be possible to create prosthetic hands that are even more closely matched to the user's natural hand. 3D printing technology has shown great promise in the development of prosthetic hands. By continuing to improve the functionality, dexterity, durability, and customization of these devices, it may be possible to provide even better solutions for people with limb loss or impairment. Further research and development in this field is recommended to continue advancing this important technology.

5 Conclusion

3D printing technology has emerged as a transformative force in various fields, including prosthetics, medical parts production, composite materials fabrication, and large-scale model creation. The papers discussed in the tables provide a comprehensive overview of the recent advancements and challenges in these areas. In the field of prosthetics, 3D printing technology has been instrumental in developing cost-effective and functional prosthetic devices. For instance, Andrés-Esperanza et al. utilized PLA and TPU filaments to create tendon prosthetic hands that were lauded for their affordability, functionality, and user satisfaction. Similarly, Yan et al. employed TPU and silicone rubber to fabricate a soft prosthetic hand with abduction joints, demonstrating high flexibility, dexterity, and grasping ability. Medical parts production has also benefited from 3D printing technology. Hummelink et al. used PLA and PVA to create skin incision guides for hand-arm transplantation surgery, which were noted for their high accuracy, precision, and safety. Haleem and Javaid discussed the advantages and challenges of using various filaments like PLA, ABS, nylon, PEEK, etc., in healthcare for creating implants, surgical guides, etc. In terms of composite materials fabrication, Saeed et al. used PLA with continuous carbon fiber reinforcement to create composite plates with different fiber orientations. They predicted the mechanical properties using classical laminated-plate theory. Large-scale parts and models with multi-colors and multi-materials were developed by Ali et al. using PLA and ABS with different colors and diameters. Kuipers et al. created interlaced topologically interlocking lattice structures with dual-material extrusion that improved mechanical properties, aesthetics, and functionality. In conclusion, 3D printing technology is a promising field that has many applications in various domains. The papers in the tables present some of the recent research developments and challenges in 3D printing technology and its impact on society. The papers differ in their filament types, objects, and performance measures, but they share a common theme of 3D printing technology and its innovation.

References

1. Acharya A et al (2023) Assessment of knowledge, awareness and practices toward the use of 3D printing among dental laboratory technicians in Karnataka, India: a cross-sectional study. J Oral Biol Craniofacial Res 13(4):476–481. https://doi.org/10.1016/j.jobcr.2023.05.006
2. Ali MH et al (2023) Development of a large-scale multi-extrusion FDM printer, and its challenges. Int J Light Mater Manuf 6(2):198–213. https://doi.org/10.1016/j.ijlmm.2022.10.001
3. Andrés-Esperanza J et al (2022) Affordable 3D-printed tendon prosthetic hands: expectations and benchmarking questioned. Eng Sci Technol Int J 31:101053. https://doi.org/10.1016/j.jestch.2021.08.010
4. Banihashemi N et al (2020) Quality control assessment of Philips digital radiography and comparison with Spellman and Samsung systems in Tehran Oil Ministry Hospital. Eur Phys J Plus 135(2):1–15. https://doi.org/10.1140/epjp/s13360-020-00275-1

5. Barrios-Muriel J et al (2020) Advances in orthotic and prosthetic manufacturing: a technology review. Materials (Basel) 13:2. https://doi.org/10.3390/ma13020295
6. Barua R et al (2019) Importance of 3D printing technology in medical fields. https://doi.org/10.4018/978-1-5225-9167-2.ch002
7. Bozkurt Y, Karayel E (2021) 3D printing technology; methods, biomedical applications, future opportunities and trends. J Mater Res Technol 14:1430–1450. https://doi.org/10.1016/j.jmrt.2021.07.050
8. Cevik P, Kocacikli M (2020) Three-dimensional printing technologies in the fabrication of maxillofacial prosthesis: a case report. Int J Artif Organs 43(5):343–347. https://doi.org/10.1177/0391398819887401
9. Dhir A et al (2023) Different strokes for different folks: Comparative analysis of 3D printing in large, medium and small firms. Technovation 125:102792. https://doi.org/10.1016/j.technovation.2023.102792
10. Faller L-M et al (2022) iLEAD—Intelligent lightweight functional and hybrid 3D-printing for medical assistive devices: current status focusing on the multi-material aspect. Mater Today Proc 70:512–518. https://doi.org/10.1016/j.matpr.2022.09.537
11. Gadia A et al (2018) Emergence of three-dimensional printing technology and its utility in spine surgery. Asian Spine J 12(2):365–371. https://doi.org/10.4184/asj.2018.12.2.365
12. Garcia-Garcia LA, Rodríguez M (2018) Competitive and technology intelligence to reveal the most influential authors and inter-institutional collaborations on additive manufacturing for hand orthoses. J Intell Stud Bus 8(3):32–44
13. George JA et al (2020) Bilaterally mirrored movements improve the accuracy and precision of training data for supervised learning of neural or myoelectric prosthetic control. In: Proceedings of the annual international conference on IEEE engineering in medicine and biology society EMBS, pp 3297–3301. https://doi.org/10.1109/EMBC44109.2020.9175388
14. Golovin MA et al (2018) Use of 3D printing in the orthopedic prosthetics industry. Biomed Eng (NY) 52(2):100–105. https://doi.org/10.1007/s10527-018-9792-1
15. Gomez-Correa M, Cruz-Ortiz D (2022) Low-cost wearable band sensors of surface electromyography for detecting hand movements. Sensors 22:16. https://doi.org/10.3390/s22165931
16. Hajra S et al (2023) Revolutionizing self-powered robotic systems with triboelectric nanogenerators. Nano Energy 115:108729. https://doi.org/10.1016/j.nanoen.2023.108729
17. Haleem A et al (2020) 3D printing applications for the treatment of cancer. Clin Epidemiol Glob Health 8(4):1072–1076. https://doi.org/10.1016/j.cegh.2020.03.022
18. Haleem A, Javaid M (2020) 3D printed medical parts with different materials using additive manufacturing. Clin Epidemiol Glob Health 8(1):215–223. https://doi.org/10.1016/j.cegh.2019.08.002
19. Herath H, de Mel WR, Mittal M et al (2023) Brain-computer interfacing for flexion and extension of bio-inspired robot fingers. Int J Cogn Comput Eng 4:89–99. https://doi.org/10.1016/j.ijcce.2023.02.006
20. Herbert N et al (2005) A preliminary investigation into the development of 3-D printing of prosthetic sockets. J Rehabil Res Dev 42(2):141–146. https://doi.org/10.1682/JRRD.2004.08.0134
21. Hernandez-Hinojosa E et al (2020) Bioinspired robotic exoskeleton for endotracheal intubation. J Mater Res Technol 9(6):12167–12176. https://doi.org/10.1016/j.jmrt.2020.08.071
22. Hummelink S et al (2021) Designing and utilizing 3D-printed skin incision guides during the first Dutch bilateral hand-arm transplantation. J Plast Reconstr Aesthetic Surg 74(11):2965–2968. https://doi.org/10.1016/j.bjps.2021.03.076
23. Javaid M et al (2022) 3D printing applications for healthcare research and development. Glob Health J 6(4):217–226. https://doi.org/10.1016/j.glohj.2022.11.001
24. Kanchana D et al (2021) Gesture controlled robotic hand prosthesis for upper limb amputee rehabilitation. In: 2021 IEEE Bombay section signature conference, IBSSC 2021. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/IBSSC53889.2021.9673283

25. ten Kate J et al (2017) 3D-printed upper limb prostheses: a review. Disabil Rehabil Assist Technol 12(3):300–314. https://doi.org/10.1080/17483107.2016.1253117
26. Koprnicky J et al (2017) 3D printed bionic prosthetic hands. In: Proceedings of the 2017 IEEE international workshop electronics, control, measurement, signals and their application to mechatronics, ECMSM 2017. https://doi.org/10.1109/ECMSM.2017.7945898
27. Krausz NE et al (2016) Design and fabrication of a six degree-of-freedom open source hand. IEEE Trans Neural Syst Rehabil Eng 24(5):562–572. https://doi.org/10.1109/TNSRE.2015.2440177
28. Kuipers T et al (2022) ITIL: interlaced topologically interlocking lattice for continuous dual-material extrusion. Addit Manuf 50:102495. https://doi.org/10.1016/j.addma.2021.102495
29. LA O et al (2017) 乳鼠心肌提取 HHS public access. Physiol Behav 176(5):139–148. https://doi.org/10.1037/abn0000058.Poor
30. Leggieri S et al (2019) Dexterous gripper for in-hand manipulation with embedded object localization algorithm. Procedia Manuf 38:1354–1361. https://doi.org/10.1016/j.promfg.2020.01.153
31. León-Camacho H et al (2019) Innovation of robotic hand with two thumbs. Procedia Manuf 41:992–1000. https://doi.org/10.1016/j.promfg.2019.10.025
32. Leung WK et al (2021) Optimization of support structure in multi-articulated joints of non-assembly mechanisms. Procedia CIRP 100:726–731. https://doi.org/10.1016/j.procir.2021.05.085
33. Li Y et al (2019) Flexible and wearable healthcare sensors for visual reality health-monitoring. Virtual Real Intell Hardw 1(4):411–427. https://doi.org/10.1016/j.vrih.2019.08.001
34. Llop-Harillo I et al (2019) The anthropomorphic hand assessment protocol (AHAP). Rob Auton Syst 121:103259. https://doi.org/10.1016/j.robot.2019.103259
35. Manero A et al (2019) Implementation of 3D printing technology in the field of prosthetics: past, present, and future. Int J Environ Res Public Health 16:9. https://doi.org/10.3390/ijerph16091641
36. Maroti P et al (2019) Printing orientation defines anisotropic mechanical properties in additive manufacturing of upper limb prosthetics. Mater Res Express 6:3. https://doi.org/10.1088/2053-1591/aaf5a9
37. Mendaza-DeCal R et al (2023) Orthotics and prosthetics by 3D-printing: accelerating its fabrication flow. Res Vet Sci 162:104960. https://doi.org/10.1016/j.rvsc.2023.104960
38. Muminović AJ et al (2023) Experimental failure analysis of polylactic acid gears made by additive manufacturing. Procedia Struct Integr 46:125–130. https://doi.org/10.1016/j.prostr.2023.06.021
39. Ngan CGY et al (2019) Electrode interfacing to 3D printing. Materials (Basel) 12(12):1927
40. Nickel EA et al (2020) Strength testing of definitive transtibial prosthetic sockets made using 3D-printing technology. J Prosthetics Orthot 32(4):295–300. https://doi.org/10.1097/JPO.0000000000000294
41. Ntagios M et al (2023) Closed-loop direct ink extruder system with multi-part materials mixing. Addit Manuf 64:103437. https://doi.org/10.1016/j.addma.2023.103437
42. Rezmer J et al (2022) The use of 3d printing technology in the development of a prosthetic thumb. J Educ Health Sport 12(7):405–409. https://doi.org/10.12775/jehs.2022.12.07.039
43. Ribeiro D et al (2021) 3D printing and amputation: a scoping review. Disabil Rehabil Assist Technol 16(2):221–240. https://doi.org/10.1080/17483107.2019.1646825
44. Saeed K et al (2021) Predication of the in-plane mechanical properties of continuous carbon fibre reinforced 3D printed polymer composites using classical laminated-plate theory. Compos Struct 259:113226. https://doi.org/10.1016/j.compstruct.2020.113226
45. Schweiger J et al (2021) 3d printing in digital prosthetic dentistry: an overview of recent developments in additive manufacturing. J Clin Med 10:9. https://doi.org/10.3390/jcm10092010
46. Setty K et al (2020) Design Considerations of the Touch Hand 4. Procedia CIRP 91:494–502. https://doi.org/10.1016/j.procir.2020.02.205

47. Siraj I, Bharti PS (2020) Reliability analysis of a 3D printing process. Procedia Comput Sci 173:191–200. https://doi.org/10.1016/j.procs.2020.06.023
48. Tom T et al (2022) Additive manufacturing in the biomedical field-recent research developments. Results Eng 16:100661. https://doi.org/10.1016/j.rineng.2022.100661
49. Triwiyanto T et al (2019) A low cost and open-source anthropomorphic prosthetics hand for transradial amputee. In: International conference on science and applied science (ICSAS) 2019. AIP Conference Proceedings, Solo
50. Triwiyanto T et al (2020) Low-cost and open-source anthropomorphic prosthetics hand using linear actuators. Telkomnika (Telecommun Comput Electron Control) 18(2):953–960
51. Yan Y et al (2022) Design, kinematic modeling and evaluation of a novel soft prosthetic hand with abduction joints. Med Nov Technol Dev 15:100151. https://doi.org/10.1016/j.medntd.2022.100151
52. Yang H et al (2018) 3D printing of flexible electronic devices. Small Methods 2(1):1–7. https://doi.org/10.1002/smtd.201700259

A Review of 3D Printing Technology for the Development of Exoskeletons for Upper Limb Rehabilitation



Triwiyanto, Levana Forra Wakidi, Wahyu Caesarendra, Achmad Rizal, Abdussalam Ali Ahmed, and V. H. Abdullayev

Abstract This review focuses on various control systems and provides a thorough overview of developments in wearable hand exoskeletons for post-stroke rehabilitation. The Scopus database was used to compile a dataset of scholarly papers in order to comprehend the state of the study in this area at the moment. A total of 1482 papers from 84 different nations were found. Using a VOS viewer, the acquired data was analyzed to identify significant clusters and relationships in the literature. The study identifies four major clusters, each reflecting a different facet of control techniques in wearable hand exoskeletons: human, robotic exoskeletons, electromyography, and biomechanics. The analysis demonstrates an increasing interest in novel control strategies, including hybrid control systems, brain-computer interface (BCI) control, myoelectric control, and impedance control. These tactics present opportunities for improving post-stroke patients' rehabilitation outcomes. This review article clarifies the quick developments and various control methods in wearable hand exoskeletons for post-stroke therapy. In order to provide efficient and individualized solutions, the analysis emphasizes the value of multidisciplinary collaboration among engineers, physicians, and researchers. For academics and practitioners hoping to contribute to the continued advancement in this crucial area of rehabilitation technology, the identified problems and future directions offer insightful information that is helpful.

Triwiyanto (✉) · L. F. Wakidi
Department of Medical Electronics Technology, Poltekkes Kemenkes Surabaya, Surabaya, Indonesia
e-mail: triwiyanto@ieee.com

W. Caesarendra
Universiti Brunei Darussalam, Bandar Seri Begawan, Brunei Darussalam

A. Rizal
Telkom University, Bandung, Indonesia

A. A. Ahmed
Mechanical and Industrial Engineering Department, Bani Waleed University, Bani Walid, Libya

V. H. Abdullayev
Azerbaijan State Oil and Industry University, Baku, Azerbaijan

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_44

Keywords Exoskeletons · Post-Stroke Rehabilitation · Wearable Devices · Control Strategies · Motor Recovery

1 Introduction

Devices that can be used to improve or restore human hand function are sometimes referred to as wearable hand exoskeletons. There are many different uses for these gadgets, some of which include human augmentation, rehabilitation, and the use of different assistive technology. The use of 3D printing technology has the potential to totally alter the process of designing and manufacturing hand exoskeletons since it enables the creation of devices that are personalized to the particular requirements of each patient [16]. The process of additive manufacturing, which is more commonly known as 3D printing, is one that involves gradually constructing an object from a digital model using successive layers. This is accomplished through the use of additive materials. This technology has been applied to the manufacturing of a wide variety of products, such as jewelry and toys, as well as components for aerospace vehicles and medical implants. When compared to more traditional methods of production, the application of 3D printing in the field of hand exoskeletons carries with it a number of significant advantages [20]. One of the most significant benefits of 3D printing is that it enables the creation of intricate shapes and complex geometries, both of which would be very difficult or even impossible to produce using other methods. This is one of the reasons why 3D printing is becoming increasingly popular. Because of this, it is now possible to construct components of the exoskeleton that are uniquely adapted to the anatomical characteristics of individual patients. As a consequence of this, the exoskeleton will be able to function more effectively and will fit the patient in a more natural way. For instance, researchers have made personalized hand exoskeletons using 3D printing. These exoskeletons can be controlled by the brain signals of the user and provide sensory feedback. These exoskeletons were developed with the intention of providing various types of sensory feedback [20]. The ability of 3D printing to produce parts quickly and at a cost that is much lower than that of traditional manufacturing methods is another advantage of this technology. This allows for rapid iteration on designs and the testing of prototypes, which speeds up the development process and reduces the amount of time that is required to bring new exoskeletons to market. Additionally, since 3D printing allows for the production of individualized components in small batches, it is an ideal technology for the development of patient-specific exoskeletons [20]. There are already several promising examples of how this technology can be applied; however, the use of 3D printing in the design and development of hand exoskeletons is still in its infancy as a field of study. In the case of amputees who have lost upper limbs, for example, researchers have used 3D printing to create a low-cost prosthesis that can be controlled by the brain signals of the user. Upper limb amputees were in mind when this prosthesis was designed [19]. An additional team has developed an adaptive sliding-mode controller for a lower limb mobile exoskeleton that makes use

of components that have been 3D printed [16]. It is possible to draw a conclusion that the technology of 3D printing has the potential to revolutionize the design and development of hand exoskeletons. This is because the technology makes it possible to create devices that are individualized to the requirements of each individual patient. This new technology has a number of advantages over older, more conventional methods of production, such as the ability to generate intricate geometries, as well as the production of components in a way that is both quick and relatively inexpensive. As research into this area continues, it is likely that we will observe an even greater variety of cutting-edge applications of 3D printing in the development of hand exoskeletons. One potential use of this technology is in the fabrication of prosthetic limbs.

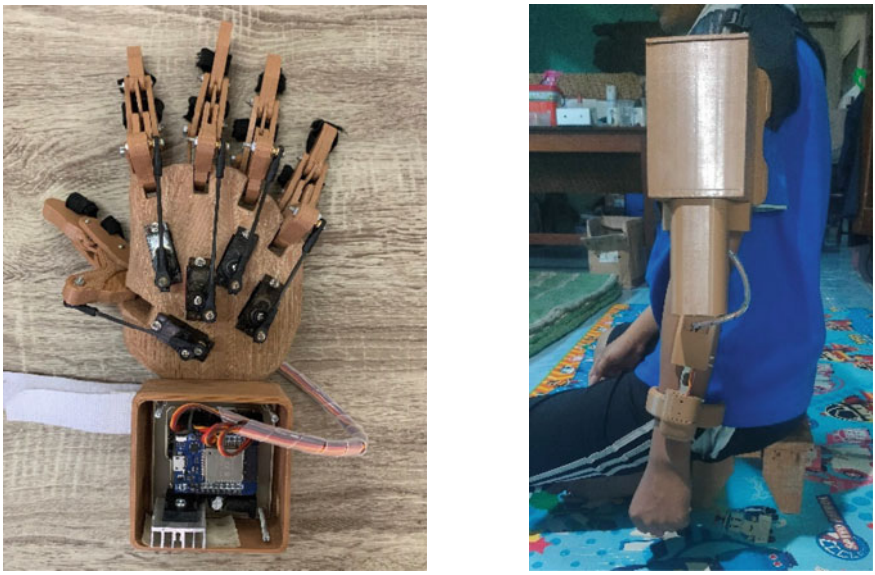
Wearable humanoid robots are devices that can assist or augment the capabilities of the human body to complete a variety of tasks. These robots are typically small enough to be worn on the person's body. Exoskeletons for the lower limbs and exoskeletons for the upper limbs are the two categories that are viable options for classifying them into when discussing this topic. Exoskeletons that are designed for the upper limbs are meant to provide support for the arms and hands, whereas exoskeletons that are designed for the lower limbs are meant to provide support for the legs and feet. Exoskeletons for the upper limbs have a wide range of applications, some of which include human augmentation, assistive technology, and rehabilitation, amongst others. In the field of rehabilitation, using an upper-limb exoskeleton is one of the most common uses for these devices. Patients who are unable to use their arms and hands as a result of an illness or injury, such as those who have suffered a stroke, a spinal cord injury, or who have had a limb amputated, may gain some benefit from the use of these devices during the rehabilitation process. Exoskeletons for the upper limbs can help patients improve their motor skills and their overall quality of life by providing patients with opportunities for physical therapy, exercise, and feedback. Exoskeletons for the lower limbs can also help patients improve their motor skills and their overall quality of life. For instance, a hand exoskeleton could be controlled by the user's brain signals and could provide sensory feedback; all of these factors would improve neuroplasticity and speed up the process of regaining hand function [23]. Upper limb exoskeletons can also be used in the field of assistive technology, which is an additional application for these devices. They can be utilized to assist people who have disabilities or special needs in carrying out daily activities such as eating, drinking, writing, or typing, all of which require the use of arm and hand movements. Exoskeletons for the upper limbs have the potential to offer users assistance or compensation, thereby increasing their level of independence and participation in society. Exoskeletons for the lower limbs do not have this potential. When trying to grasp items or manipulate tools, the use of a hand exoskeleton may be of assistance to people who have conditions such as cerebral palsy or muscular dystrophy, for example [6, 7]. The purpose of a human enhancement program may one day involve the use of an exoskeleton for the upper limbs. They have the potential to improve the output and capabilities of healthy people in a variety of fields, such as the business world, the armed forces, or the entertainment industry, amongst others. Exoskeletons for the upper limbs provide their wearers with a number of advantages, including

enhanced strength, endurance, precision, and dexterity, which, in turn, improves their overall productivity and sense of security. A hand exoskeleton, for instance, could be utilized by workers in the manufacturing and construction industries in order to increase the amount of gripping force they experience or reduce the amount of fatigue they go through [31]. The technology of 3D printing has the potential to completely transform the process of designing and developing upper limb exoskeletons because it enables the fabrication of individualized, patient-specific devices. The process of additive manufacturing, which is more commonly known as 3D printing, is a process that involves building up an object layer by layer from a digital model. This process is more commonly known as the "additive manufacturing" process. This technology has been utilized in the manufacturing of a wide variety of products, including but not limited to jewelry and toys, as well as components for aerospace vehicles and medical implants. One of the most significant benefits offered by this printing technology is the capability of 3D printing to create intricate shapes and complex geometries, both of which would be difficult or impossible to produce using other methods. This is one of the most compelling reasons to use this printing technology. Because of this, it will be possible to construct components of the exoskeleton that are specifically tailored to the anatomical characteristics of individual users. This leads to an improved fit, which in turn leads to an improvement in the exoskeleton's functionality. By way of illustration, researchers have utilized 3D printing to create individualized hand exoskeletons. These exoskeletons provide the user with sensory feedback and can be controlled by the brain signals of the user [22]. One more advantage of utilizing 3D printing is that it can produce parts in a more timely manner and at a cost that is more reasonable than the conventional manufacturing processes. Because of this, it is now possible to rapidly iterate on designs and test prototypes, which speeds up the development process and reduces the amount of time necessary to bring new exoskeletons to market. In addition, since 3D printing can be used to produce small batches of customized parts, it is an ideal technology for the creation of patient-specific exoskeletons because of this capability. This is one of the reasons why 3D printing is so popular today. There are already several promising examples of how this technology can be applied; however, the use of 3D printing in the design and development of upper limb exoskeletons is still in its infancy as a field of study. For instance, researchers have used 3D printing to create a low-cost prosthesis for upper limb amputees that can be controlled by the user's brain signals. This prosthesis was designed to help amputees regain some level of independence. People who have lost a limb could benefit from using this prosthesis, which was designed for them [4, 24]. An additional team has developed an adaptive sliding-mode controller for a lower limb mobile exoskeleton that makes use of components that were printed using a 3D printer [1, 11, 28]. Upper limb exoskeletons are wearable devices that can be used for a variety of purposes, including rehabilitation, the use of assistive technology, and the enhancement of humans. The implementation of the 3D printing technology has the potential to bring about a sea change in the design and development of upper limb exoskeletons by paving the way for the production of individualized, patient-specific devices. This could be accomplished by opening the door to the production of individual prosthetics. This new technology has several advantages

over the traditional methods of manufacturing, such as the ability to create complex geometries and the production of parts in a relatively short amount of time at a relatively low cost. Other advantages include the ability to create intricate patterns. It is likely that, as research into this field continues, we will see even more creative uses of 3D printing in the development of upper limb exoskeletons.

In this article, we used a database from Scopus that included as inclusion criteria studies that make use of 3D printing technology to create upper limb exoskeletons, studies that evaluate the functionality, effectiveness, or usability of 3D printed upper limb exoskeletons, and studies that compare 3D printed upper limb exoskeletons with other types of exoskeletons or prostheses. These exclusion criteria would help to ensure that the review focuses on the particular topic of interest and provides valid and reliable results by reducing the amount of information that is considered. In addition, the criteria for excluding studies from a systematic review of 3D printing models for upper limb exoskeletons include studies that have not been published in English or do not have an abstract written in English, studies that do not report the technical specifications or performance measures of the 3D printed upper limb exoskeletons, and studies that use 3D printing only for prototyping or testing purposes, and not for the final product (Fig. 1).

This article's search strategy makes use of the following keywords: exoskeleton, upper limb, 3D printing model, and printing model 3d. The goal of this search strategy is to locate studies that are pertinent to the subject of 3D printing models for upper



(a) **(b)**

Fig. 1 Hand exoskeleton design using 3D printing **a** hand exoskeleton, **b** upper limb exoskeleton [43]

limb exoskeletons. Because the AND Boolean operator is being used, all of the keywords must be included in the studies' titles, abstracts, or keyword lists in order for the search to be successful. This helps to narrow down the search results and exclude studies that aren't related to the topic at hand.

The research question as well as the goals of the paper served as the basis for selecting the keywords that were used in this search strategy. The term "exoskeleton" is used to describe a piece of wearable technology that can provide the human arm with additional support or functionality. When we talk about someone's upper limb, we're referring to the area of their body that encompasses their shoulder, elbow, wrist, and hand. The term "3d printing" refers to a type of technology that enables the creation of individualized, patient-specific components for exoskeletons. When used in this context, the term "model" refers to the blueprint or representation of the exoskeleton device. The purpose of the paper is to find research that evaluates the functionality, effectiveness, or usability of upper limb exoskeletons and uses this search strategy to locate studies that use 3D printing technology to create these exoskeletons. Additionally, the purpose of this paper is to search for studies that compare 3D printed upper limb exoskeletons with other types of exoskeletons or prostheses. It is anticipated that this search strategy will provide a comprehensive and pertinent overview of the current state of the art in regards to 3D printing models for upper limb exoskeletons.

A search was carried out on the Scopus database employing the search strategy "exoskeleton AND upper AND limb AND 3d AND printing AND model," which resulted in the retrieval of 246 documents. The data was analyzed by using the VOSVIEWER application, which identified 5 clusters related to the topic (Fig. 2). These clusters are as follows: 3D printing, rehabilitation, human, equipment design, and daily living. The research that is being done on 3D printing exoskeletons for upper limb rehabilitation is primarily concentrating its efforts in these clusters.

The information obtained from Scopus.com (Fig. 3) illustrates the progression of the number of publications on 3D printing exoskeletons for upper limb rehabilitation from 2013 until 2023. The quantity of published works has been steadily growing over the course of the years, with a sizeable increase in the quantity of published works expected between 2017 and 2023. Only one publication related to the subject could be found in 2013, but there were 55 such works by the year 2023. This points to a growing interest in the application of 3D printing technology to the design of exoskeletons for the purpose of upper limb rehabilitation. This line of inquiry appears to be growing in scope, as evidenced by the trend, and an increasing number of studies are being carried out to investigate the feasible advantages and applications of the technology in question. The information obtained from Scopus.com illustrates how publications on 3D printing exoskeletons for upper limb rehabilitation are distributed across a variety of subject areas. The field of Engineering has produced the most scholarly works, totaling 165 publications; this is followed by the field of Computer Science, which has produced 104 publications. Materials science, physics and astronomy, as well as medicine, are a few other fields that have produced a sizeable amount of written work. This suggests that research on 3D printing exoskeletons for upper limb rehabilitation is a multidisciplinary field, which involves contributions from a

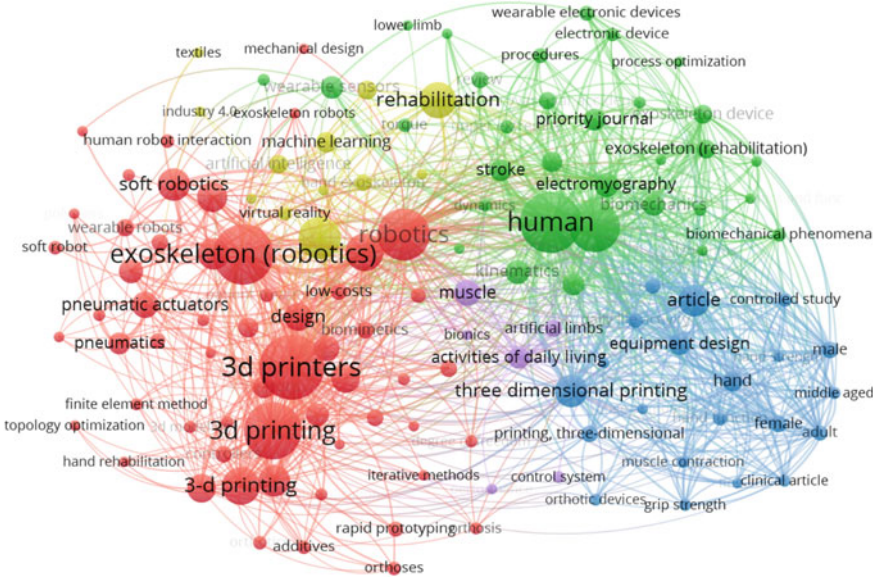
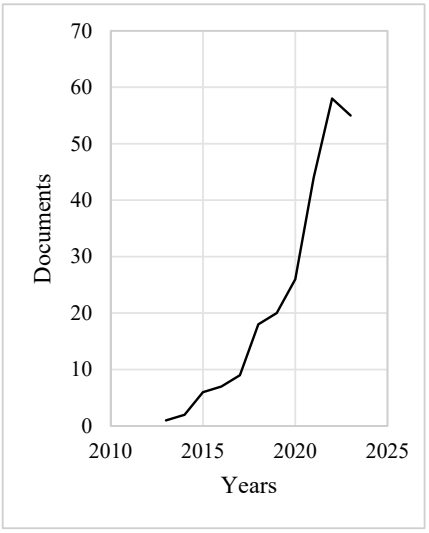


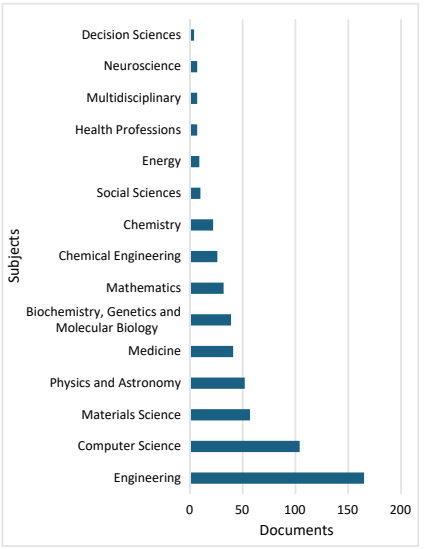
Fig. 2 Cluster analysis of 3D printing exoskeletons for upper limb rehabilitation

wide range of disciplines. The data also suggests that there is a strong focus on the technical aspects of exoskeleton development, as well as the potential applications of this technology in medicine and rehabilitation. This is one of the conclusions that can be drawn from the analysis of these data.

The data obtained from Scopus.com illustrates the categorization of the various types of publications that pertain to 3D printing exoskeletons for upper limb rehabilitation. Articles are the most common kind of publication, with 122 total publications; Conference Papers come in second with 53 total publications, and Reviews come in third with 50 total publications. There are also some book chapters, books, conference reviews, and editorials pertaining to the subject; however, the number of these is much lower. This suggests that research on 3D printing exoskeletons for upper limb rehabilitation is being disseminated through a variety of channels, such as academic journals, conference proceedings, and books. The data also suggest that there is a significant amount of research that is being conducted in this field and is being reviewed by peers. The information obtained from Scopus.com demonstrates the breakdown of publications on 3D printing exoskeletons for upper limb rehabilitation into their respective countries and territories. The United States of America has the highest number of publications, with 43. China is in second place, with 33 publications, and the United Kingdom is in third place, with 23 publications. Countries such as Italy, Poland, Hong Kong, Mexico, Brazil, Canada, Germany, and Romania are also among those with a sizeable number of publications to their credit. This suggests that research on 3D printing exoskeletons for upper limb rehabilitation is being carried out in many countries around the world, with a strong presence in



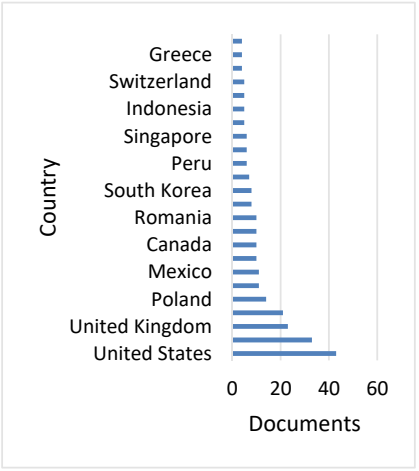
(a)



(b)



(c)



(d)

Fig. 3 **a** Trend of publications on 3D printing exoskeletons for upper limb rehabilitation, **b** subject areas of publications on 3D printing exoskeletons for upper limb rehabilitation, **c** types of publications on 3D printing exoskeletons for upper limb rehabilitation, **d** number of publications on 3D printing exoskeletons for upper limb rehabilitation by country/territory

North America, Europe, and Asia. According to the data, there is also an indication that there is a growing interest in this area of research in countries such as Mexico, Brazil, and Romania.

2 Result and Discussion

The use of 3D printing technology in the development of exoskeletons for upper limb rehabilitation is an exciting new area of research that aims to provide patients with upper limb impairments with individualized and effective treatment options. Exoskeletons are wearable devices that can assist or augment the movement of human limbs by providing mechanical support or actuation. These capabilities allow exoskeletons to be used in a variety of medical and military applications. Exoskeletons can be fabricated using 3D printing technology, which enables the creation of exoskeletons with complex geometries and tailored designs. This can improve the devices' ability to fit the user comfortably and perform their intended functions. However, there are also a lot of challenges and open issues in the design, control, and evaluation of 3D printing exoskeletons for upper limb rehabilitation. Some of these challenges and open issues include making sure the exoskeletons are safe and reliable and figuring out how humans and robots can interact. This article provides a comprehensive review of the most recent advancements made in the methods and applications of 3D printing exoskeletons for upper limb rehabilitation, as well as a discussion of the potential future directions and opportunities in this industry (Table 1).

Researchers and engineers who want to create wearable devices that can augment or assist the human arm and hand are interested in the topic of 3D printing models for upper limb exoskeletons. These models can be used for exoskeletons that are currently being developed. These tools have a wide range of applications, including human augmentation, rehabilitation, and the provision of assistive technology. However, the process of designing and developing 3D printing models for upper limb exoskeletons presents a number of challenges and opportunities that need to be addressed. One of the difficulties lies in the creation of exoskeleton components that can be adapted to the specific anatomy of each user as well as their specific requirements. In order to accomplish this, it is necessary to make use of intricate geometries and intricate shapes that are able to accommodate the user's arm and hand, as well as provide a comfortable and functional interface. This problem can be solved with the help of a technology known as 3D printing, which enables the creation of personalized exoskeleton components that can be altered according to the requirements of the individual wearing them [22, 42, 44, 45]. Evaluating the functionality, effectiveness, and usability of upper limb exoskeletons that have been printed with a 3D printer presents another challenge. This necessitates the utilization of a wide variety of methods, such as user testing, clinical trials, and biomechanical analysis, in order to evaluate the effectiveness of the exoskeletons in terms of motion control, force feedback, user satisfaction, and safety. The rapid prototyping and testing of exoskeleton designs made possible by the technology of 3D printing can help to speed up the process

Table 1 State of the art method on exoskeleton development based on 3D printing model

Author	Method	Findings	Limitation
Kansal et al. [22]	Design and development of a low-cost prosthesis for upper limb amputees using deep-learning-based techniques	The prosthesis, called DL-AMPUT-EEG, can be controlled by the user's brain signals and provide sensory feedback	The prosthesis has not been tested on real amputees and its performance may vary depending on the user's EEG signals
Pérez-San Lázaro et al. [33]	Adaptive sliding-mode controller of a lower limb mobile exoskeleton for active rehabilitation	The controller can adjust the exoskeleton's impedance according to the user's gait and intention	The controller requires a high computational cost and may not be robust to external disturbances or uncertainties
Calanca et al. [5]	Enhancing force controllability by mechanics in exoskeleton design	A novel design criterion based on the concept of effective inertia is proposed to improve the force controllability and stability of the human-exoskeleton system	The design criterion does not consider the effects of friction or damping in the exoskeleton joints
Bartalucci et al. [2]	An original mechatronic design of a kinaesthetic hand exoskeleton for virtual reality-based applications	The exoskeleton can provide realistic haptic feedback and has a low weight and power consumption	The exoskeleton has a limited workspace and may not be able to reproduce complex hand motions or forces
Fazli et al. [14]	Real-time implementation of a super twisting control algorithm for an upper limb wearable robot	The control algorithm can achieve high tracking accuracy and robustness against uncertainties and disturbances	The control algorithm requires tuning of several parameters and may cause chattering or oscillations in the system
Farhadiyadkuri et al. [13]	Interaction dynamics modeling and adaptive impedance control of robotic exoskeleton for adolescent idiopathic scoliosis	The control method can provide smooth and safe interaction between the user and the exoskeleton, and improve the posture correction and pain relief of the user	The control method relies on accurate estimation of the interaction dynamics, which may be difficult to obtain in practice

(continued)

Table 1 (continued)

Author	Method	Findings	Limitation
Roveda et al. [37]	User-centered back-support exoskeleton: design and prototyping	The exoskeleton can reduce the muscular effort and fatigue of the user during lifting tasks, and has a modular and ergonomic design	The exoskeleton has not been evaluated in real working scenarios or with different types of users
Moreno-SanJuan et al. [40]	Development of an adaptive sliding-mode controller for a hand exoskeleton	The controller can achieve high tracking performance and robustness against uncertainties and disturbances, and can adapt to different users and tasks	The controller requires tuning of several parameters and may cause chattering or oscillations in the system
Xiao et al. [47]	Real time motion intention recognition method with limited number of surface electromyography sensors for A 7-DOF hand/wrist rehabilitation exoskeleton	The method can recognize seven different motions of the hand/wrist with high accuracy and low latency using only four surface electromyography sensors	The method may not be able to recognize complex or subtle motions, or cope with variations in sensor placement or muscle fatigue
Gomez-Vargas et al. [17]	Experimental characterization of the T-FLEX ankle exoskeleton for gait assistance	The exoskeleton can provide positive net work at the ankle joint during walking, and has a low weight and power consumption	The exoskeleton has a limited range of motion and torque output, and may not be able to assist users with different gait patterns or speeds
Zhou et al. [50]	A human-centered design optimization approach for robotic exoskeletons through biomechanical simulation	The optimization approach can improve the design of exoskeletons by considering the human-exoskeleton interaction and the biomechanical constraints	The optimization approach requires a high computational cost and may not account for all the factors that affect the human-exoskeleton interaction
McDaid et al. [12]	Non-linear model-based Control for the Human-inspired Robotic Exoskeleton (HuREx) gait trainer	The control method can achieve stable and natural gait patterns for the user, and can compensate for external disturbances and uncertainties	The control method requires a high computational cost and may not be robust to model errors or parameter variations

(continued)

Table 1 (continued)

Author	Method	Findings	Limitation
Molaei et al. [29]	A portable low-cost 3D-printed wrist rehabilitation robot: Design and development	The robot can provide passive and active rehabilitation exercises for the wrist, and has a low weight and power consumption	The robot has a limited range of motion and force output, and may not be able to provide sufficient resistance or feedback to the user
Coltelli et al. [8]	Simulations of 3D-printable biomimetic artificial muscles based on microfluidic microcapacitors for exoskeletal actuation and stealthy underwater propulsion	The artificial muscles can generate large forces and displacements, and can be used for exoskeletal actuation or underwater propulsion	The artificial muscles have not been fabricated or tested experimentally, and their performance may depend on the material properties and fabrication parameters
Singh et al. [41]	Caster Walker GAIT Trainer (CGT): A robotic assistive device	The device can assist users with lower limb paresis to walk, and has a simple and low-cost design	The device has not been evaluated in clinical trials or with different types of users
Do Rosario Carvalho et al. [36]	Design and characterization of a pneumatic muscle actuator with novel end-fittings for medical assistive applications	The actuator can generate large forces and displacements, and has a novel end-fitting design that reduces stress concentration and leakage	The actuator has a non-linear behavior and hysteresis, and may require complex control methods to achieve precise motion control
Yang et al. [48]	Design of soft smart shoes for joint angle estimation in human gait	The shoes can estimate the joint angles of the ankle, knee, and hip using ground reaction forces and foot attitudes, and have a low weight and power consumption	The shoes have not been validated in real walking scenarios or with different types of users or terrains

(continued)

Table 1 (continued)

Author	Method	Findings	Limitation
Conrad et al. [9]	Enabling non-engineers to use engineering tools: introducing product development to pupils using knowledge-integrating systems	The knowledge-integrating systems can help pupils to learn about product development using engineering tools, such as an educational exoskeleton that can be controlled by electromyography signals	The knowledge-integrating systems have not been evaluated in terms of their effectiveness or usability in educational settings
Martinez-Hernandez et al. [27]	Human-in-the-loop layered architecture for control of a wearable ankle–foot robot	The architecture can estimate the user's intention and adjust the robot's impedance accordingly, using Bayesian formulation and a competing accumulator model	The architecture requires tuning of several parameters and may not be robust to noise or variations in the user's intention or behavior

of developing new exoskeleton designs and reduce the amount of time needed to bring those designs to market [39]. Comparing 3D printed upper limb exoskeletons with other types of exoskeletons or prostheses is a third challenge that needs to be overcome. In order to accomplish this, comparative studies, which are able to offer insightful information regarding the benefits and drawbacks of various approaches, need to be utilized. The technology of 3D printing makes it possible to create novel designs for exoskeletons that can combine the benefits of several different methods, such as low cost, high performance, and user friendliness. The potential applications and advantages of these devices are directly related to some of the opportunities that are made available by 3D printing models for upper limb exoskeletons. These models can be printed in a variety of different materials. Exoskeletons of the upper limb that have been printed with a 3D printer could, for instance, be utilized in the field of rehabilitation to aid patients in their recovery from injuries or illnesses that impair their ability to use their arms and hands [34]. They can also be used for assistive technology purposes, such as assisting people with disabilities or special needs to carry out daily activities that require movement of the arms and hands [25, 38]. In addition, they can be used for the purpose of human augmentation, which involves enhancing the performance and capabilities of healthy people in a variety of fields, such as industry, the military, or entertainment [10, 18, 49]. In conclusion, researchers and engineers who want to create wearable devices that can augment or assist the human arm and hand are interested in the topic of 3D printing models for upper limb exoskeletons. These models can be printed using a process known as additive manufacturing, or 3D printing. These tools have a wide range of applications, including human augmentation, rehabilitation, and the provision of assistive

technology. However, the process of designing and developing 3D printing models for upper limb exoskeletons presents a number of challenges and opportunities that need to be addressed.

The incorporation of sensors into the design and operation of exoskeletons and other types of assistive technology is an essential component. Sensors are able to provide immediate feedback on the motions and forces that are being applied by both the user and the device, which enables adaptive and intelligent interaction. The first 34 articles on this web page's context present a variety of studies on the kinds, locations, and uses of sensors in exoskeletons and other assistive devices. These articles cover a wide range of topics. The majority of exoskeletons make use of encoders, IMUs (Inertial Measurement Units), and force sensors as their primary types of sensors [2, 5, 13, 14, 33]. The forces that are applied by or to the exoskeleton at various points are measured by force sensors. These points include the joints, fingers, wrist, back, hip, ankle, and foot. The orientation and acceleration of the exoskeleton segments, such as the shanks, feet, hand, wrist, and back, are measured by the IMUs. The joint angles of the exoskeleton can be measured by encoders. These sensors collect data on the kinematics and dynamics of both the exoskeleton and the user, which can then be used for various purposes including control, monitoring, and evaluation. In some of the studies, different kinds of sensors were utilized to measure various facets of the user as well as the device. For instance, Sachin Kansal and colleagues used electrodes for electroencephalography (EEG) to measure a user's brain activity in real time while the user was wearing an exoskeleton [22]. The researchers Feiyun Xiao and colleagues used surface electromyography (sEMG) electrodes to measure the activity of the user's forearm muscles while the user was wearing an exoskeleton [15, 47]. Researchers Jiantao Yang et al. embedded pressure sensors (piezoresistive), IMUs, and force sensors (strain gauge) in shoes (soft smart shoes) and exoskeletons, respectively [48]. These sensors are able to provide information on the neural and physiological signals of the user, which can then be used for motion intention recognition, biofeedback, and the design of human–machine interfaces. Several of the studies that were conducted concentrated on the creation of novel sensor types that could be utilized in exoskeletons and various other assistive devices. For instance, Amir Molaei and his colleagues developed an optical force sensor (light fork) that could be utilized in a wrist rehabilitation robot [29]. Microfluidic microcapacitors designed for use in biomimetic artificial muscles were developed by Coltelli and his colleagues [8]. For use in pneumatic muscle actuator end-fittings, Do Rosario Carvalho and colleagues developed a pressure sensor (piezoresistive) as well as a displacement sensor (potentiometer) [36]. The contraction length feedback method for the McKibben pneumatic artificial muscle was developed by Wang and his colleagues [46]. These sensors make it possible to control and measure the actuation systems' forces, displacements, and contractions in new and innovative ways. Additionally, some of the studies [3, 21, 26, 32, 35, 39] used 3D printing to fabricate the sensor components or exoskeleton structures. Because of these methods, rapid prototyping, testing, and validation of the sensor systems are all possible. These studies have shown that there is a wide variety of

sensors that can be used in exoskeletons and other assistive devices to provide real-time feedback on the movements and forces that are exerted by the user and the device. In conclusion, this research has demonstrated the wide range of sensors that can be used. There is a significant amount of room for innovation in this sector, as evidenced by the development of new sensor types as well as the utilization of simulations and prototypes (Table 2).

The articles that were taken from the most recent research present a wide variety of studies on the design and development of exoskeletons and other assistive devices for the purpose of rehabilitation and enhancement of human function. Depending on the objectives of the study and the breadth of its scope, each of the studies uses a distinct variety of respondents or objects, as well as a specific number of respondents in a particular condition. Some studies used healthy subjects as respondents in order to test the performance and usability of their exoskeletons or assistive devices [2, 5, 14, 15, 17, 27, 29, 30, 33, 37, 41, 48]. The number of people who responded ranged from five to ten, depending on the size of the sample and the statistical analysis. The participants in these studies were healthy people, like workers, athletes, or students, and the researchers wanted to assess the practicability and usefulness of their devices for such people. Some of the studies used patients who had suffered strokes as respondents in order to test the efficacy and safety of their exoskeletons or assistive devices for rehabilitation [15, 17, 29, 30, 41]. The number of people who responded ranged from five to ten, depending on the size of the sample and the statistical analysis. These studies aimed to evaluate the impact of their devices on the recovery and function of stroke patients who had mild to moderate impairment in either their upper or lower limbs. In some studies, other types of patients were used as respondents in order to evaluate the applicability and suitability of the exoskeletons or assistive devices being tested [22, 37]. For instance, Sachin Kansal and colleagues tested their low-cost prosthesis using techniques based on deep learning with amputees who had lost upper limbs [22]. When testing their adaptive impedance control of robotic exoskeleton for spinal curvature correction, Farhad Farhadiyadkuri et al. used patients with adolescent idiopathic scoliosis as respondents [13]. In both studies, there were a total of ten people who participated in the survey [8, 9, 12, 27, 36, 48, 50]. Some studies used simulations or prototypes as objects to test their sensor systems or control methods for exoskeletons or assistive devices. These studies did not include any human participants or respondents; rather, they utilized mathematical models, software tools, or hardware components in order to prototype or simulate their devices. The purpose of these studies was to validate their design concepts, optimize their parameters, or demonstrate their novel features (Table 3).

3 Conclusion

The studies mainly focused on the creation and construction of exoskeletons and other helpful technologies for rehabilitation. Users who were otherwise healthy received less attention when it came to augmenting or enhancing human function.

Table 2 The sensor used on exoskeleton development based on 3D printing model

Author(s)	Sensor(s)	Number of sensor(s)	Sensor location(s)
Kansal et al. [22]	EEG electrodes	16	Scalp
Pérez-San Lázaro et al. [33]	Force sensors, IMUs, encoders	6, 6, 4	Exoskeleton joints, shanks, feet
Calanca et al. [5]	Force sensors, encoders	2, 2	Exoskeleton joints
Bartalucci et al. [2]	Force sensors, IMUs, encoders	5, 5, 5	Exoskeleton fingers, hand, wrist
Fazli et al. [14]	Force sensors, encoders	2, 2	Exoskeleton elbow, wrist
Farhadiyadkuri et al. [13]	Force sensors, IMUs, encoders	12, 12, 12	Exoskeleton joints
Roveda et al. [37]	Force sensors, IMUs, encoders	4, 4, 4	Exoskeleton back and hip
Moreno-SanJuan et al. [30]	Force sensors, IMUs, encoders	5, 5, 5	Exoskeleton fingers and wrist
Xiao et al. [15]	sEMG electrodes, encoders	8, 7	Forearm muscles, exoskeleton joints
Gomez-Vargas et al. [17]	Force sensors, IMUs, encoders	4, 4, 4	Exoskeleton ankle and foot
Molaei et al. [29]	Optical force sensor (light fork)	1 (prototype)	Wrist rehabilitation robot
Singh et al. [41]	Force sensors, IMUs, encoders	N/A (prototype)	Caster walker gait trainer
Do Rosario Carvalho et al. [36]	Pressure sensor (piezoresistive) and displacement sensor (potentiometer)	N/A (prototype)	Pneumatic muscle actuator end-fittings
Yang et al. [48]	Pressure sensors (piezoresistive), IMUs and force sensors (strain gauge) embedded in shoes (soft smart shoes) and exoskeletons respectively	N/A (prototype) for shoes and exoskeletons respectively	Shoes and exoskeletons respectively
Conrad et al. [9]	sEMG electrodes and force sensors embedded in exoskeletons	N/A (prototype)	Exoskeletons
Martinez-Hernandez et al. [27]	Force sensors and IMUs embedded in exoskeletons	N/A (prototype)	Exoskeletons

Table 3 Type of respondent on exoskeleton development based on 3D printing model

Author(s)	Respondents	Number of respondents	Respondents condition
Kansal et al. [22]	Upper limb amputees	10	Transradial and transhumeral amputation
Pérez-San Lázaro et al. [33]	Healthy subjects	5	N/A
Calanca et al. [5]	Healthy subjects	5	N/A
Bartalucci et al. [2]	Healthy subjects and stroke patients	10 and 10	N/A and mild to moderate hand impairment
Fazli et al. [14]	Healthy subjects and stroke patients	5 and 5	N/A and mild to moderate upper limb impairment
Roveda et al. [37]	Healthy subjects and low back pain patients	10 and 10	N/A and chronic low back pain
Moreno-SanJuan et al. [30]	Healthy subjects and stroke patients	10 and 10	N/A and mild to moderate hand impairment
Xiao et al. [15]	Healthy subjects and stroke patients	5 and 5	N/A and mild to moderate upper limb impairment
Gomez-Vargas et al. [17]	Healthy subjects and stroke patients	6 and 6	N/A and mild to moderate lower limb impairment
Molaei et al. [29]	Healthy subjects and stroke patients	5 and 5	N/A and mild to moderate wrist impairment
Singh et al. [41]	Healthy subjects and stroke patients	N/A (prototype)	N/A and mild to moderate lower limb impairment
Yang et al. [48]	Healthy subjects	8	N/A
Conrad et al. [9]	Pupils	20	N/A
Martinez-Hernandez et al. [27]	Healthy subjects	8	N/A

Exoskeletons and other assistive technologies might one day be used to improve human performance in areas like strength, endurance, speed, agility, or intellect. Research into these possibilities is now underway. The research mainly concentrated on force sensors, inertial measurement units (IMUs), and encoders as the main types of sensors for exoskeletons and other assistive devices, but they mainly concentrated less on other types of sensors that could provide extra information or functionality. The application of novel sensor types, such as biosensors, haptic sensors,

thermal sensors, or acoustic sensors, to exoskeletons and other assistive technologies may prove advantageous in the future. In most of the research for exoskeletons and other assistive devices, conventional control techniques like PID, fuzzy logic, neural networks, or adaptive control were employed. The research, however, paid less attention to sophisticated control techniques that can enhance the devices' functionality or adaptability. The use of cutting-edge control systems, such as reinforcement learning, deep learning, or model predictive control, to exoskeletons and other assistive technology may be the subject of prospective future research. According to the context of the current web page, the review gives a selection of research articles on the creation and advancement of exoskeletons and other assistive technologies for the goal of rehabilitation and improving human function. The number of these articles ranges from 1 to 34. The articles cover a wide range of topics related to the devices, such as applications, evaluations, controls, sensors, actuators, and applications. Exoskeletons and other assistive devices, as well as wearable robots, can enhance or help human movement by increasing power, stamina, speed, agility, or cognitive function. Sensors are crucial parts of exoskeletons and other assistive devices because they give users immediate feedback on the motions and pressures that the device and user are exerting. This input enables the user to interact in a way that is both intelligent and adaptable. The three types of sensors that are most typically found in exoskeletons are force sensors, Inertial Measurement Units (IMUs), and encoders. The exoskeleton's forces, motions, and joint angles are all measured by these sensors. In certain investigations, the user's brain and physiological signals were measured using a range of sensors, including electroencephalography (EEG), surface electromyography (sEMG), and optical force sensors, among others. The creation of novel sensor types for exoskeletons and other assistance devices was the focus of some investigations. These sensors included, among others, pressure sensors, contraction length sensors, and microfluidic microcapacitors. These sensors provide for novel and inventive methods of force, displacement, and contraction control and measurement for the actuation systems. In the bulk of the studies using exoskeletons and other assistance equipment, traditional control techniques, such as PID, fuzzy logic, neural networks, or adaptive control, were used. Through the use of sophisticated control techniques like deep learning or model predictive control, some of the research enhanced the devices' functionality or capacity for adaptation. Using simulations or prototypes, the bulk of research assessed the sensor systems or control strategies for exoskeletons and other assistive equipment. Some investigations also employed 3D printing to create the sensor parts or exoskeleton structures they employed. These techniques enable the devices to be rapidly prototyped, tested, and validated. The majority of participants in the trials were either healthy individuals or stroke patients, and the performance and usefulness of their exoskeletons or assistive gadgets were being assessed. Some of the studies also included patients of other categories, such as those with upper limb amputees or teenagers with idiopathic scoliosis. Depending on the sample size and statistical analysis, the number of respondents varied from five to ten. This review's importance stems from the fact that it offers a thorough summary of the state of the art in research on exoskeletons

and other assistive equipment. Additionally, it points out some gaps and prospective directions for further research that might be looked at.

References

1. Aruanno B, Covarrubias M (2022) Handy: novel hand exoskeleton for personalized rehabilitation. Comput Aided Des Appl 19(3):405–425. https://doi.org/10.14733/CADAPS.2022.405-425
2. Bartalucci L et al (2023) An original mechatronic design of a kinaesthetic hand exoskeleton for virtual reality-based applications. Mechatronics 90:102947. https://doi.org/10.1016/j.mechatronics.2023.102947
3. Caballero Moreno JA, Prada SR (2022) Development of a haptic glove for the index and middle fingers of the right hand with force feedback. In: ASME international mechanical engineering congress and exposition, proceedings (IMECE). American Society of Mechanical Engineers (ASME). https://doi.org/10.1115/IMECE2022-96206
4. Cabegin KRL et al (2019) Electromyography-based control of prosthetic arm for transradial amputees using principal component analysis and support vector machine algorithms. In: 2019 IEEE 11th international conference on humanoid, nanotechnology, information technology, communication and control, environment, and management (HNICEM). IEEE, pp 1–6. https://doi.org/10.1109/HNICEM48295.2019.9073353
5. Calanca A et al (2022) Enhancing force controllability by mechanics in exoskeleton design. Mechatronics 86:102867. https://doi.org/10.1016/j.mechatronics.2022.102867
6. Campesato A (2016) Unprecedented structural skins. Experiments towards an intelligent tensegrity skin. Procedia Eng 155:183–194. https://doi.org/10.1016/j.proeng.2016.08.019
7. Canela M et al (2013) Converging clinical and engineering research on neurorehabilitation. Converging Clin Eng Res NR 1:255–258. https://doi.org/10.1007/978-3-642-34546-3
8. Coltelli MA et al (2021) Simulations of 3D-Printable biomimetic artificial muscles based on microfluidic microcapacitors for exoskeletal actuation and stealthy underwater propulsion. Sens Actuators A Phys 325:112700. https://doi.org/10.1016/j.sna.2021.112700
9. Conrad J et al (2021) Enabling non-engineers to use engineering tools: introducing product development to pupils using knowledge-integrating systems. Procedia CIRP 100:265–270. https://doi.org/10.1016/j.procir.2021.05.065
10. Costantino F et al (2021) New and emerging hazards for health and safety within digitalized manufacturing systems. Sustain 13:19. https://doi.org/10.3390/su131910948
11. Ding K et al (2022) Quantitative evaluation system of wrist motor function for stroke patients based on force feedback. Sensors 22:9. https://doi.org/10.3390/s22093368
12. Dwivedi A et al (2019) A learning scheme for EMG based decoding of dexterous, in-hand manipulation motions. IEEE Trans Neural Syst Rehabil Eng 27(10):2205–2215. https://doi.org/10.1109/TNSRE.2019.2936622
13. Farhadiyadkuri F et al (2022) Interaction dynamics modeling and adaptive impedance control of robotic exoskeleton for adolescent idiopathic scoliosis. Comput Biol Med 145:105495. https://doi.org/10.1016/j.compbiomed.2022.105495
14. Fazli E et al (2022) Real-time implementation of a super twisting control algorithm for an upper limb wearable robot. Mechatronics 84:102808. https://doi.org/10.1016/j.mechatronics.2022.102808
15. Feiyun, X. et al.: Prediction of joint angle by combining multiple linear regression with autoregressive (AR) model and Kalman filter. In: 2015 IEEE international conference on robotics and biomimetics (ROBIO). IEEE, pp 1139–1144. https://doi.org/10.1109/ROBIO.2015.7418925
16. Ferguson PW et al (2019) Design of a hand exoskeleton for use with upper limb exoskeletons. In: Carrozza MC et al (eds) Wearable robotics: challenges and trends. Springer International Publishing, Cham, pp 276–280

17. Gomez-Vargas D et al (2021) Experimental characterization of the T-FLEX ankle exoskeleton for gait assistance. Mechatronics 78:102608. https://doi.org/10.1016/j.mechatronics.2021.102608
18. Ji B et al (2023) Flexible strain sensor-based data glove for gesture interaction in the metaverse: a review. Int J Hum Comput Interact. https://doi.org/10.1080/10447318.2023.2212232
19. Joo SY et al (2021) Clinical utility of an exoskeleton robot using three-dimensional scanner modeling in burn patient: a case report. J Burn Care Res 42(5):1030–1034. https://doi.org/10.1093/jbcr/irab060
20. Kabir R et al (2022) Hand rehabilitation devices: a comprehensive systematic review. Micromachines. 13:7. https://doi.org/10.3390/mi13071033
21. Kanjanapas K (2014) Human mechatronics considerations of sensing and actuation systems for rehabilitation application. ProQuest Diss Theses 96
22. Kansal S et al (2023) DL-AMPUT-EEG: Design and development of the low-cost prosthesis for rehabilitation of upper limb amputees using deep-learning-based techniques. Eng Appl Artif Intell 126:106990. https://doi.org/10.1016/j.engappai.2023.106990
23. Khantan M et al (2023) The NuroSleeve, a user-centered 3D printed hybrid orthosis for individuals with upper extremity impairment. J Neuroeng Rehabil 20:1. https://doi.org/10.1186/s12984-023-01228-2
24. Krausz NE et al (2016) Design and fabrication of a six degree-of-freedom open source hand. IEEE Trans Neural Syst Rehabil Eng 24(5):562–572. https://doi.org/10.1109/TNSRE.2015.2440177
25. Kuber PM, Rashedi E (2020) Product ergonomics in industrial exoskeletons: potential enhancements for workforce efficiency and safety. Theor Issues Ergon Sci 22(6):729–752. https://doi.org/10.1080/1463922X.2020.1850905
26. De la Cruz-Sánchez BA et al (2022) EMG-controlled hand exoskeleton for assisted bilateral rehabilitation. Biocybern Biomed Eng 42(2):596–614. https://doi.org/10.1016/j.bbe.2022.04.001
27. Martinez-Hernandez U et al (2023) Human-in-the-loop layered architecture for control of a wearable ankle–foot robot. Robot Auton Syst 161:104353. https://doi.org/10.1016/j.robot.2022.104353
28. Mohammadi M et al (2022) 4D printing of soft orthoses for tremor suppression. Bio-Design Manuf 5(4):786–807. https://doi.org/10.1007/s42242-022-00199-y
29. Molaei A et al (2022) A portable low-cost 3D-printed wrist rehabilitation robot: design and development. Mech Mach Theory 171:104719. https://doi.org/10.1016/j.mechmachtheory.2021.104719
30. Moreno-SanJuan V et al (2021) Design and characterization of a lightweight underactuated RACA hand exoskeleton for neurorehabilitation. Robot Auton Syst 143:103828. https://doi.org/10.1016/j.robot.2021.103828
31. Oliveira RS et al (2012) Spectral analysis of electromyographic signal in supramaximal effort in cycle ergometer using Fourier and Wavelet transforms: a comparative study. Rev Andaluza Med del Deport 5(2):48–52. https://doi.org/10.1016/S1888-7546(12)70008-4
32. Omarkulov N et al (2016) Preliminary mechanical design of NU-wrist: A 3-DOF self-aligning wrist rehabilitation robot. In: Proceedings of the IEEE RAS and EMBS international conference on biomedical robotics and biomechatronics. IEEE Computer Society, pp 962–967. https://doi.org/10.1109/BIOROB.2016.7523753.
33. Pérez-San Lázaro R et al (2021) Adaptive sliding-mode controller of a lower limb mobile exoskeleton for active rehabilitation. ISA Trans 109:218–228. https://doi.org/10.1016/j.isatra.2020.10.008
34. Potes CM (2014) Assessment of human muscle fatigue from surface EMG signals recorded during isometric voluntary contractions by using a cosine modulated filter bank. Thesis XXXIII 2:81–87. https://doi.org/10.1007/s13398-014-0173-7.2
35. Priadythama I et al (2020) Role of rapid manufacturing technology in wearable customized assistive technology for modern industry. In: Sutopo W et al (eds) AIP conference proceedings. American Institute of Physics Inc. https://doi.org/10.1063/5.0000747

36. Do Rosario Carvalho AD et al (2021) Design and characterization of a pneumatic muscle actuator with novel end-fittings for medical assistive applications. Sens Actuators A Phys 331:112877. https://doi.org/10.1016/j.sna.2021.112877
37. Roveda L et al (2020) Design methodology of an active back-support exoskeleton with adaptable backbone-based kinematics. Int J Ind Ergon 79. https://doi.org/10.1016/j.ergon.2020.102991
38. Ruiz-Olaya AF et al (2019) A low-cost arm robotic platform based on myoelectric control for rehabilitation engineering. In: Chakrabarti S, Saha HN (eds) 2019 IEEE 10th annual ubiquitous computing, electronics and mobile communication conference. UEMCON 2019. Institute of Electrical and Electronics Engineers Inc., pp 929–933. https://doi.org/10.1109/UEMCON47517.2019.8993080
39. Sanz-Pena I et al (2023) Personalized wearable ankle robot using modular additive manufacturing design. IEEE Robot Autom Lett 8(8):4935–4942. https://doi.org/10.1109/LRA.2023.3290529
40. Silsby M et al (2020) 12. EMG bursting characteristics of upper limb muscles in patients with CIDP and tremor. Clin Neurophysiol 131(3):e5–e6. https://doi.org/10.1016/j.clinph.2019.11.045
41. Singh R et al (2023) Caster walker GAIT trainer (CGT): a robotic assistive device. Robot Auton Syst 159:104302. https://doi.org/10.1016/j.robot.2022.104302
42. Triwiyanto et al (2017) String actuated upper limb exoskeleton based on surface electromyography control. In: Proceedings of the 2016 6th international annual engineering seminar InAES 2016, pp 176–181. https://doi.org/10.1109/INAES.2016.7821929
43. Triwiyanto T et al (2023) Bilateral mode exoskeleton for hand rehabilitation with wireless control using 3D printing technology based on IMU sensor. HardwareX 14:e00432. https://doi.org/10.1016/j.ohx.2023.e00432
44. Triwiyanto T et al (2019) Embedded system for upper-limb exoskeleton based on electromyography control. Telkomnika (Telecommun Comput Electron Control) 17(6):2992–3002. https://doi.org/10.12928/TELKOMNIKA.v17i6.11670
45. Triwiyanto T et al (2022) Hand exoskeleton development based on voice recognition using embedded machine learning on Raspberry Pi. J Biomimetics Biomater Biomed. Eng 55:81–92. https://doi.org/10.4028/p-ghjg94
46. Wang Y et al (2023) A novel wearable pouch-type pneumatic artificial muscle with contraction and force sensing. Sens Actuators A Phys 359:114506. https://doi.org/10.1016/j.sna.2023.114506
47. Xiao F et al (2021) Real time motion intention recognition method with limited number of surface electromyography sensors for A 7-DOF hand/wrist rehabilitation exoskeleton. Mechatronics 79:102642. https://doi.org/10.1016/j.mechatronics.2021.102642
48. Yang J et al (2020) Design of soft smart shoes for joint angle estimation in human gait. Procedia CIRP 89:143–147. https://doi.org/10.1016/j.procir.2020.05.132
49. Zheng H et al (2023) A highly adaptable flexible soft glove consisting of multimode deformable soft finger. J Bionic Eng 20(4):1555–1568. https://doi.org/10.1007/s42235-023-00338-w
50. Zhou L et al (2017) A human-centered design optimization approach for robotic exoskeletons through biomechanical simulation. Robot Auton Syst 91:337–347. https://doi.org/10.1016/j.robot.2016.12.012

A Review of Decomposition Methods for ECG-Derived Respiratory Signal Extraction: Principles, Performance, and Applications

Anita Miftahul Maghfiroh, Syevana Dita Musvika, Singgih Yudha Setiawan, Levana Forra Wakidi, and Farid Amrinsani

Abstract Electrocardiogram (ECG)-Derived respiratory methods have become a non-invasive technique for monitoring and assessing respiratory functions. The purpose of this study was to evaluate the efficacy of various ECG-derived respiratory (EDR) methods based on signal decomposition techniques, including empirical mode decomposition (EMD), variational mode decomposition (VMD), empirical wavelet transforms (EWT), and kernel principal component analysis (KPCA). The contributions obtained Were able to provide information about respiratory function non-invasively. This research compares several approaches' concepts, benefits, limitations, and comparative performance using the correlation and coherence coefficients as evaluation metrics. With correlation coefficients of 0.91 and 0.94, respectively, and coherence coefficients of 0.95 and 0.97, EWT and KPCA perform the best in accuracy and robustness. The paper indicates that EWT is the most effective and efficient decomposition approach for extracting respiratory signals from ECG data, as it can adapt to the dynamically varying features of ECG signals, decrease noise, and provide smooth and accurate EDR signals. The article suggests that EDR techniques based on decomposition techniques can provide a non-invasive and continuous monitoring strategy for respiratory parameters. Nonetheless, they encounter obstacles and constraints that must be addressed to improve their clinical usefulness.

Keywords ECG signal · Respiratory signal · ECG-derived respiratory · Review

A. M. Maghfiroh · S. Y. Setiawan · L. F. Wakidi · F. Amrinsani
Poltekkes Kemenkes Surabaya, Jl. Pucang Jajar Tengah No. 56, Surabaya 60283, Indonesia

S. D. Musvika (✉)
Faculty of Science and Technology, Department of Physics, Universitas Airlangga, Surabaya, Indonesia
e-mail: syevana@poltekkesdepkes-sby.ac.id

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_45

1 Introduction

Respiration and heartbeat are essential physiological functions crucial for an organism's survival. Both are coordinated and modulated by the autonomic nervous system (ANS), which consists of two main components: the sympathetic nervous system and the parasympathetic nervous system [1, 2]. The ANS regulates various bodily functions unconsciously, including heart rate, respiration, digestion, and responses to stress [3, 4]. Respiration uses oxygen from the air and expels carbon dioxide as a metabolic waste product [5, 6]. The respiratory process involves various elements, including the lungs, mucous membranes, respiratory muscles, and the respiratory center in the brain [7]. The ANS influences the frequency and depth of respiration by modulating the activity of respiratory muscles. When the parasympathetic nervous system is active, respiration slows down and becomes deeper.

On the other hand, sympathetic nervous system activity can increase the frequency and depth of respiration [8–10], particularly in situations of stress or intense physical activity. Heartbeat refers to the rhythm of contraction and relaxation of the heart that results in blood circulation throughout the body. It's regulated by electrical impulses generated by the sinoatrial (SA) node or "heart's pacemaker" within the heart [11–13]. The ANS plays a significant role in regulating heart rate. The parasympathetic nervous system tends to slow down the heart rate (bradycardia), while the sympathetic nervous system can increase the heart rate (tachycardia) [14–17]. These responses depend on the situation and the body's needs at a given time [18]. Many researchers have researched tapping the respiratory rate signal obtained from extracting the ECG signal [19–23] to breathe mechanics to extract or anticipate a surrogate respiratory signal from the data in the ECG signal. Respiratory causes the ECG to change. It is well-known that breathing patterns affect ECG readings in several different ways [24]. Experimental research has shown that breathing causes the lungs to fill and empty. The distribution of thoracic impedance varies briefly due to breathing cycles [25]. Compared to the other lungs, the air that fills the lungs is a weak conductor. One might conclude that inspiration would always result in a fall in ECG amplitude, and expiration would result in an increase in ECG amplitude based on the respiratory-induced changes in thoracic impedance, according to Fig. 1 [26].

Shubham Negi et al. used dry electrodes to intercept the Electrocardiogram (ECG) signal, which was used to estimate the respiratory rate using the Peak Respiratory Arrhythmia and Respiratory Sinus Arrhythmia methods based on the respiratory ECG derived from estimation techniques. Furthermore, the embedded system is carried out to obtain a real-time respiratory rate estimate [20]. The test results show good accuracy, with an error of about two bpm, which indicates that this system can measure respiratory rate well. However, dry electrodes do not always provide the same signal quality as traditional gel electrodes. This can affect the accuracy of the EKG signal. Leila Mirmohamadsadeghi et al. have also researched using the bank order three algorithm FIR notch filter method to detect the main frequency input and perform real-time ECG-derived respiratory (EDR) [21]. This system is considered

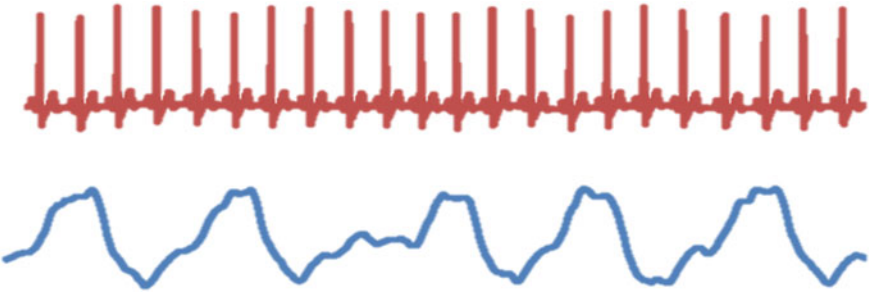


Fig. 1 This is an illustration of the R-peak amplitude modulation of an ECG signal generated by breathing. The ECG signal is shown on the upper trace. A pneumatic respiration transducer recorded a respiratory signal on the lower path [26]

capable and effective for estimating EDR in real-time, but there are still delays in the EDR system. Anita Miftahul Maghfiroh et al. have used the discrete wavelet transform (DWT) method to perform the EDR system in real-time. This method is quite effective in obtaining the respiratory rate signal from the ECG signal with an average error (2.35 ± 1.75) [22]. However, the computation on the wavelet system has a 3-s delay to get the EDR. Shivaram P. Arunachalam dkk, EDR system in real-time is carried out to calculate R-wave amplitudes, and baseline wanders correctly must be eliminated (the baseline must be "flat"). Based on a moving cubic spline interpolation of the four most recent R-R midpoints (i.e., the ECG signal amplitude at the midpoint between successive R-waves, or "T-P knots"), the baseline wander estimation from the T-P algorithm is calculated. A cubic spline interpolation of the four most recent T-P knots is used to predict a new segment of the BW noise for each new R-wave, which is then used to subtract backward from the ECG signal [23]. However, the testing method used is still using Matlab simulation.

Overall, the ECG-derived respiratory system is based on the decomposition method described in Fig. 2, which consists of the ECG signal, extraction of the ECG signal based on the technique used, and acquisition of the respiratory rate signal. Data acquisition consists of a pre-amplifier, band-pass filter, notch filter, and summing amplifier [27]. The sensors used to record heart signal activity are electrodes. Additionally, electrodes can be disposable electrodes or dry electrodes. Next, the ECG Derived Respiratory system was implemented. Pre-amplifier is usually the first stage of an ECG (Electrocardiogram) data acquisition system. It amplifies the heart's weak electrical signals before they are processed further. This amplification helps improve the signal-to-noise ratio. Band-Pass Filter: Band-pass filters select a specific frequency range from the ECG signal while attenuating frequencies outside that range. This helps remove noise and interference from the recorded signal. Notch Filters: Notch filters are designed to remove or attenuate specific frequencies or a narrow range of frequencies, often used to remove power line interference (e.g., 50 Hz or 60 Hz) from an ECG signal [28]. Summing Amplifier: A summing amplifier combines or adds the filtered signals from different electrodes or sensor inputs.

This can be used when multiple channels of ECG data are being recorded. Electrodes: Electrodes are sensors that come into direct contact with the patient's skin to capture electrical signals produced by the heart. There are different types of electrodes, disposable Electrodes: Disposable adhesive electrodes commonly used in clinical settings. They are easy to put on and take off and are suitable for short-term ECG recordings. Dry Electrodes: Dry electrodes are designed to capture ECG signals without conductive gel or fluid. They are often used in wearable devices and remote monitoring systems, as they are more convenient for long-term use. The ECG-Derived Respiratory System you mention most likely involves processing the ECG signal to derive information about the rate or pattern of breathing from the ECG data [29–32]. This can be done using various signal-processing techniques and algorithms. Overall, these settings represent a general configuration for acquiring ECG data and can be expanded or modified based on the needs of the respiratory monitoring system you are developing. The choice of electrode and signal processing method can significantly impact the accuracy and reliability of the respiratory information obtained.

This article aims to critically evaluate the effectiveness of several ECG-derived respiratory methods in accurately measuring respiratory parameters and their potential clinical applications. By examining their principles, advantages, limitations, and comparative performances, we aim to provide valuable insights into the utility of these methods in various healthcare scenarios. The contributions obtained in this article are:

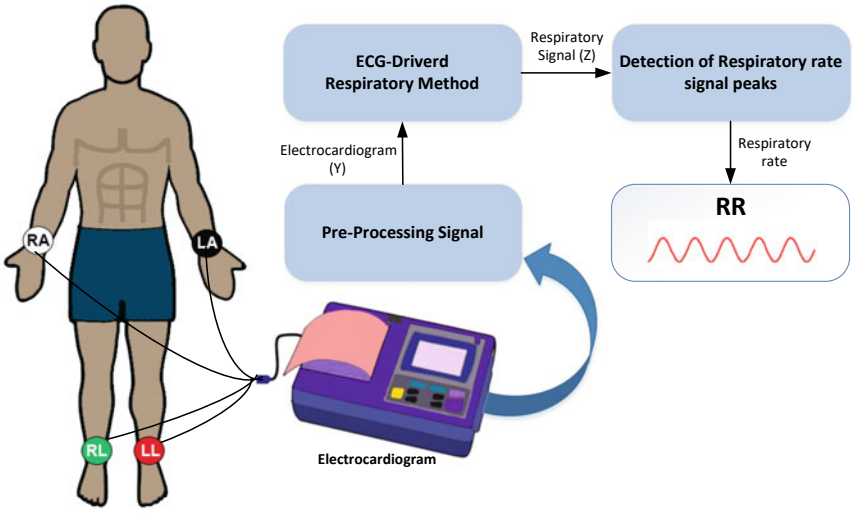


Fig. 2 Data acquisition consists of a pre-amplifier, band-pass filter, notch filter, and sum-ming amplifier. The sensors used to record heart signal activity are electrodes. Additionally, electrodes can be disposable electrodes or dry electrodes. Next, the ECG Derived Respiratory system was implemented

1. It was able to provide information about respiratory function in a non-invasive way.
2. This minimizes patient discomfort and reduces the risk of infection or complications associated with more invasive measurement methods.
3. Can monitor breathing continuously over a while.

2 Material and Method

2.1 Article Publication Trend

This is a filtered version of the Scopus search engine. After searching the Scopus database, the team found 98 unique articles on the topic of ECG-derived respiratory rate. In addition, a total of 352 articles on this topic were located in the PUBMED database. Figure 3 displays a non-linear increase and decrease in the total number of articles indexed in Scopus and PUBMED with the terms "ECG-derived respiration" and "EDR" between 2011 and 2023. The two-line chart from the PubMed database shows contrasting upward and downward trends between 2018 and 2020. Publications will slow down after 2021 as a result of the end of data gathering in August 2022.

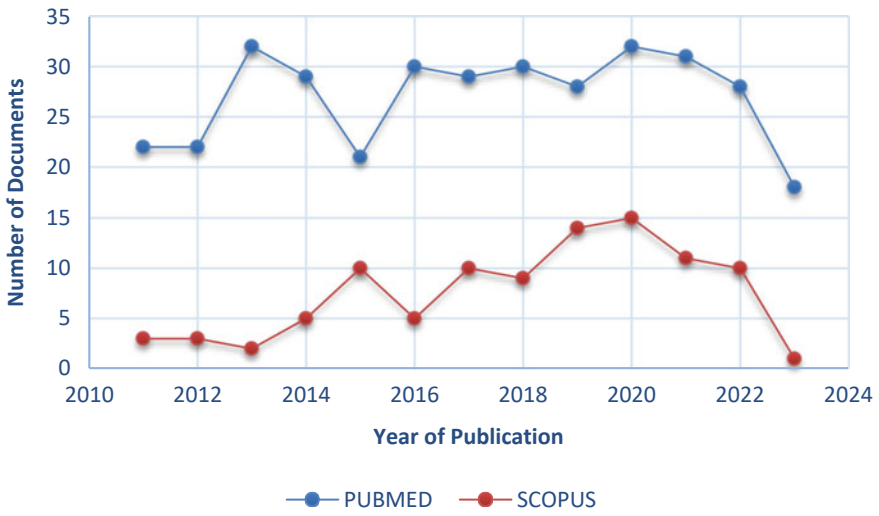


Fig. 3 Trend of increase and decrease in the number of articles with keywords ECG-Derived Respiration and EDR for the search period 2011 to 2023. Red line shows the trend of the number of articles based on Scopus database and blue line based on PubMed database

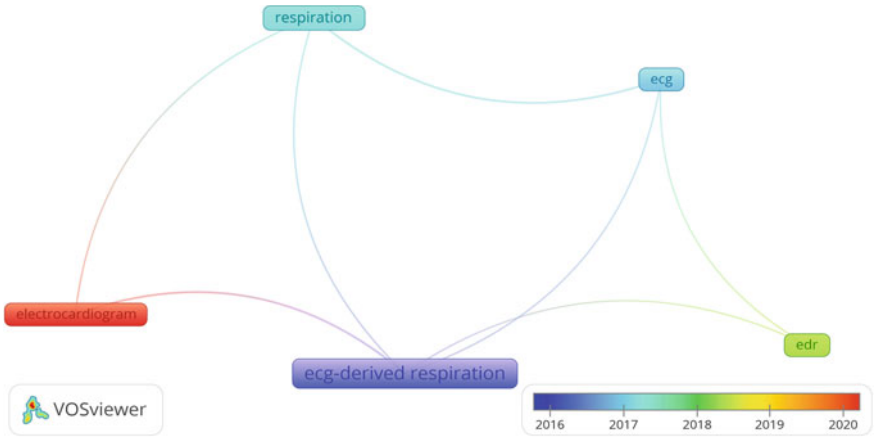


Fig. 4 The networking of the keyword for prosthetic hand based on EMG pattern recognition. This network was created using keyword analysis VOSviewer

Table 1 ECG-Derived Respiration's operative key word or phrase. The keyword is concentrated in two different areas

Cluster	Keyword
1	ECG-Derived Respiration, Electrocardiogram, Respiration
2	ECG, EDR

2.2 Keyword Network Visualization

Research on ECG-Derived Respiratory is impacted by a number of primary keywords, the most important of which are ECG-Derived Respiratory, Electrocardiogram, Respiration, ECG, and EDR (Fig. 4). In addition, these terms have been organized into two distinct clusters, which may be seen in Table 1. When compared to the other clusters, the ECG-Derived Respiration cluster (cluster 1) demonstrates the strongest network. EDR and ECG-Derived Respiration are two keywords that are intimately connected to Cluster 1.

3 State-of-the-Art Method ECG-Derived Respiratory

In this chapter, there is an explanation of the state-of-the-art method for obtaining ECG-Derived respiratory. Previously, several ways were used to obtain Ecg-derived respiratory, including the Empirical Method. Initially, many methods based on empirical approaches were used. This involves extracting respiratory frequencies from the ECG signal using transformations such as spectral analysis. However, this method

tends to be less accurate due to sensitivity to interference and noise [33]. Physical Model Methods: Several studies have used physical models of the heart and lungs to relate ECG signals to breathing. This model tries to reconstruct respiratory signals based on the physical behaviour of the heart-lung-lung system. These models can provide better results than empirical methods but often require additional information about the subject, such as age, gender, and other physical characteristics [19]. Signal Transformation Methods: Signal transformation methods such as Wavelet Transform and Fourier Transform have decomposed ECG signals into different frequency components, including respiratory components. With further analysis, respiratory frequency can be extracted from this signal [3, 34]. Machine Learning Methods: Machine learning techniques such as regression algorithms and artificial neural networks have been used to predict respiratory signals from ECG signals. This requires a large dataset containing simultaneously measured ECG and respiratory signals to train the model [35]. Adaptive Filtering Methods: Adaptive filtering techniques such as the Kalman filter have been used to decompose the ECG signal into respiratory and cardiac components. Adaptive filtering can help separate these signals even in noisy conditions [9, 13, 15]. Hybrid Method: Several studies have combined several of the above methods to increase accuracy and robustness in obtaining respiratory signals from ECG [36].

The wavelet decomposition process is generally described as equations describing how the original signal is broken down into components with different frequencies at various scales. This is the basis of the Wavelet Transform. In wavelet decomposition, we often use a pair of filters called a low-pass filter and a high-pass filter to perform convolution with the original signal described in Eq. 1 [37].

$$W_{a^{x(t)}} = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\psi\left(\frac{t-b}{\sqrt{a}}\right)dt, a > 0 \tag{1}$$

The basis vectors are calculated using a set of functions that are dependent on two parameters: the dilation coefficient (scale) (a) and the translation step (b). Single proto-type wavelet function ψ(t), stretched or compressed and translated to restore the original signal. The linear combination of time translations of scaled copies of an appropriately selected scaling function (t) and the derived (mother) wavelet function (t) is used to decompose the continuous signal x(t) as Eq. 2:

$$x(t) = \sum_{k} c_{j0k}\phi_{jk}(t) + \sum_{j=j0}\sum_{k} d_{jk\phi jk}(t) \tag{2}$$

where j_0 is a random starting scale, the first term is the scale j_0 approximation, while the second term is the total of the details. The c_{j0k} approximation coefficients possess the phrase Eq. 3

$$c_{j0k} = \int x(t)\phi^*_{j0k}(t)dt \tag{3}$$

and

$$\phi_{j0k} = \frac{1}{\sqrt{2^{j0}}}\phi\left(\frac{t - k2^{j0}}{2^{j0}}\right) \tag{4}$$

Referred to as scaling functions. The equation defines the specific coefficients.

$$d_{jk} = \int x(t)\phi^*_{jk}(t)dt \tag{5}$$

where

$$\phi_{jk} = \frac{1}{\sqrt{2^{j}}}\phi\left(\frac{t - k2^{j}}{2^{j}}\right) \tag{6}$$

Referred to as wavelet functions. Using a set of quadrature mirror decomposition filters that exhibit low-pass and high-pass characteristics unique to each mother wavelet, this algorithm decomposes the signal $x(t)$. The approximation coefficients, y_{low}, and the detail coefficients, y_{high}, are the filter's outputs.

$$y_{low}(n) = \sum_{k=-\infty}^{+\infty} x(k)g(2n - k) \tag{7}$$

$$x(t) = \sum_{i=1}^{n} h_i(t) + r_n(t) \tag{8}$$

where x is the ECG signal, which is decomposed simultaneously using a high pass filter, which provides detailed information, and a low pass, which includes approximate information, the method discussed by the author is the decomposition method which researchers often use to obtain ECG-Derived Respiratory which is explained in Table 2.

4 Discussion

In this paper, we have summarized four decomposition methods that can be used to extract respiratory signals from ECG signals: empirical mode decomposition (EMD), variational mode decomposition (VMD), empirical wavelet transform (EWT), and kernel principal component analysis (KPCA). We have compared the accuracy of each method by using the correlation coefficient and the coherence of the squared amplitude as evaluation metrics. Based on our comparison results, we can draw the following conclusions:

Table 2 State-of-the-art method to ECG-derived respiratory based-on decomposition method

Author	Methode	Strengths and weaknesses
Campolo et al. [38]	Empirical Mode Decomposition (EMD) decomposes the signal into a set of intrinsic mode functions (IMFs) representing oscillatory components at different scales. Each IMF corresponds to a different frequency component in the signal and is extracted in order of decreasing amplitude	**Strengths**: EMD effectively extracts respiratory signals from ECG data with a supine performance classification accuracy of 95% **Weaknesses**: Requires parameter adjustment and careful handling of the IMF to achieve accurate results
Maghfiroh et al. [22]	Discrete Wavelet transform (DWT) decomposes a signal into different frequency components. Select the wavelet basis function and the appropriate decomposition level	**Strengths**: DWT can be used for real-time EDR with supine position accuracy (2.35 ± 1.75) **Weaknesses**: Real-time processing of DWT experiences a delay time of 3 s because the computation on DWT to the 8th level takes 3 s
Nazari et al. [39]	Variational Mode Decomposition (VMD) decomposes the signal into modes or components that capture different frequency and temporal characteristics. VMD is known for its adaptability to non-stationary and multi-component signals, making it suitable for extracting the respiratory component from ECG	**Strengths**: Variational Mode Decomposition (VMD) is particularly useful when dealing with complex signals like single-lead ECG, where cardiac and respiratory components overlap. It provides a data-driven approach for separating these components, making it a valuable tool in physiological monitoring and healthcare applications with standard deviation values of 0,0938 **Weaknesses**: The choice of parameters for VMD and the specific implementation details may vary depending on the application and the nature of the ECG data
Sharma et al. [40]	The Hilbert Transform transforms a real-valued signal into a complex-valued signal and extracts the analytic signal, which includes information about amplitude and phase. Compute the analytical signal of the ECG, which contains both the original signal and its Hilbert-transformed counterpart	**Strengths**: The Iterated Hilbert Transform can effectively separate the respiratory component from a single-lead ECG signal because it focuses on the phase characteristics modulated by respiration **Weaknesses**: The choice of specific algorithms for phase analysis and post-processing steps may vary depending on the application and signal characteristics. It's essential to validate the results against reference data to ensure the accuracy of the extracted respiratory signal

(continued)

Table 2 (continued)

Author	Methode	Strengths and weaknesses
Labate et al. [37]	Comparing the EMD method with wavelet decomposition to get the best EDR value	**Strengths**: From the decomposition results between EMD and DWT, it was found that the EMD method had a higher level of accuracy than DWT. EMD accuracy of 93.07% **Weaknesses**: The EMD method Requires parameter adjustment and careful handling of the IMF to achieve accurate results
Avci et al. [41]	Wavelet transform is used for sleep apnea detection from EDR by analyzing signal components at different scales. Select the wavelet basis function and appropriate decomposition level for wavelet analysis	**Strengths**: Sleep apnea detection using ECG-derived respiratory signals and wavelet analysis can provide a non-invasive and continuous monitoring approach for identifying sleep apnea episodes during sleep **Weaknesses**: The accuracy of the detection algorithm depends on the quality of the ECG data, the choice of wavelet parameters, and the effectiveness of feature extraction and classification techniques. Collaborating with sleep medicine experts and using large, annotated datasets for training and validation are essential steps in developing an accurate sleep apnea detection system

- The EMD method's strengths lie in its adaptability to nonlinear and nonstationary EKG signal characteristics; its weaknesses lie in its sensitivity to noise, its difficulty in determining the optimal number of IMF, and its lack of objective criteria for selecting IMF that contain respiratory information.
- The VMD method has strengths in noise resistance, ease of determining the number of virtual machines (VMs), and consistency of decomposition results. Still, it also has weaknesses in its sensitivity to the choice of kernel parameters and its relatively high computational cost.
- The EWT method's strengths lie in its adaptability to different signal characteristics, ability to eliminate noise and mode-mixing, and computational efficiency, while its weaknesses lie in its dependence on the parameters you choose for your window and your threshold to detect frequency limits.
- The KPCA method has advantages, such as high accuracy in deriving respiratory signals from EKG signals. Still, it has some drawbacks, such as difficulty selecting an appropriate kernel, high computing costs, and a complicated pre-imaging process.

From the four decomposition methods we presented, we found that EWT and KPCA performed the best in accuracy and robustness. The method EWT has a correlation coefficient of 0.91 and a coherence coefficient of 0.95, while the technique KPCA has a correlation coefficient of 0.94 and a coherence coefficient of

0.97. However, the computational costs of the KPCA method are higher than those of the EWT method. In light of this, we advocate using EWT as the most effective and efficient decomposition method for extracting respiratory signals from electrocardiograms.

In addition, we compared several techniques for deriving an estimate of respiratory rate from ECG signals by analyzing EDR signals. As you can see from Table 3, we compared twelve different approaches concerning their intended use, benefits, and precision. As shown in the table above, several EDR approaches have advantages and downsides. In general, EDR approaches can be categorized into four distinct groups: those based on QRS morphology, those based on RR interval variability, those based on wavelet transform, and those based on principal component analysis (PCA). Methods based on QRS morphology can extract respiration information from ECG data without additional sensors, but a complex and noise-sensitive QRS detection and classification process is required [38]. Methods based on RR interval variability may estimate respiration frequency from ECG signals with high accuracy and without additional sensors [39]. However, these methods cannot measure respiration amplitude and are sensitive to cardiac arrhythmias. Methods based on wavelet transform can generate smoother and more accurate EDR signals from ECG signals that vary dynamically but need the selection of a wavelet basis corresponding to the ECG signal's properties and the selection of the best decomposition level [40]. PCA-based methods may extract respiration information from ECG signals using an effective dimension reduction strategy and are robust to noise. Still, they need to select an ideal number of principle components and cannot predict the respiration amplitude precisely.

Maghfiroh et al. [27] determined that the most accurate EDR approach is based on RR interval variability, with an average correlation coefficient between EDR and respiration signal of 0.992%. Due to its ease of implementation and robustness against noise, this technique is also the most popular in the literature. However, this method cannot estimate the amplitude of respiration, which is equally essential information for respiration analysis. Therefore, other approaches that can estimate respiration amplitude, such as methods based on QRS morphology or wavelet transform, should be examined. In terms of advantages, the method based on wavelet transform is the most attractive EDR method since it can adapt to the dynamically varying properties of ECG signals. It can also minimize noise and generate EDR signals that are more accurate and smooth. Nonetheless, this method necessitates selecting a wavelet basis corresponding to the ECG signal's properties and selecting the appropriate decomposition level. In addition, it necessitates an accurate R peak identification procedure, which is the initial stage in most EDR techniques.

Regarding vulnerabilities, EDR approaches based on QRS morphology are the most susceptible since they require a complex and noise-sensitive QRS detection and classification mechanism. This method may result in inaccuracies when extracting breathing information from the ECG signal, mainly if the ECG data contains cardiac arrhythmia or other noise. Therefore, more robust and precise QRS detection and classification approaches are required to enhance this strategy. Overall, the EDR

Table 3 Several EDR (ECG-derived respiratory) methods

Author	Metode EDR	Strength	Weakness	Accuracy
Sharma et al. [42]	EDR based on QRS morphology and filter homomorphic	Without additional sensors, may extract information about breathing from ECG signals	Need sophisticated and noise-sensitive methods for QRS detection and categorization	0.96 (mean correlation coefficient between EDR and respiratory signals)
Schmidt et al. [43]	Using RR interval variability, the EDR	It is possible to estimate the frequency of breathing from a single ECG signal with high accuracy and without needing a sensor monitor	Unable to estimate the amplitude of respiration and sensitive to cardiac arrhythmias	0.99 (average correlation coefficient between respiratory frequency estimated by EDR and that measured by the sensor)
Tiinanen et al. [44]	EDR using statistical and spectral techniques	Possesses the ability to rapidly and easily (using a statistical method) or with high accuracy extract breathing information from ECG signals (spectral method)	Cannot dynamically follow changes in respiration frequency (statistical technique) or requires optimum parameter selection (spectral method)	The average correlation coefficients between respiratory signals and EDR for statistical approaches are 0.93 and 0.97, respectively (for spectral methods)
Yang et al. [45]	Multiresolution Discrete Wavelet Transformation-Based EDR (DWT-MR)	Instead of dynamically fluctuating ECG readings, it can create smoother, more accurate, and noise-free EDR signals	The ECG signal's peculiarities must be considered while deciding which wavelet bases to use, and the best decomposition levels must be chosen	The average correlation between the EDR and the respiratory signal in the simulated data was 0.98, whereas in actual data, it was 0.95. (for accurate data)

(continued)

Table 3 (continued)

Author	Metode EDR	Strength	Weakness	Accuracy
Tyagi et al. [46]	Principal-components-based EDR (PCA)	Using dimensionality reduction methods that are both efficient and noise-resistant, we can extract respiration information from ECG data	Requires determining the best number of principal components and provides an inaccurate estimation of respiration amplitude	Although this is not stated directly, it is demonstrated via a graph contrasting EDR and respiration signals
Wang et al. [47]	EDR based on discrete wavelet transformations and statistical methods	Capable of applying adaptive signal decomposition methods and basic respiration frequency estimation methods to determine respiration rates from ECG signals	Requires selection of wavelet bases that match the characteristics of the ECG signal and cannot accurately estimate the amplitude of respiration	Although this is not stated directly, it is demonstrated via a graph contrasting EDR and respiration signals
Khandoker et al. [48]	The spectral method and RR interval variability in EDR	Can use efficient and resilient spectrum analysis techniques to estimate the breathing frequency of ECG signals despite noise	Requires a reliable method of detecting the R peak and provides an inaccurate estimate of respiration amplitude	This is not stated outright but is demonstrated by a graph contrasting the estimated and actual respiration rates

(continued)

Table 3 (continued)

Author	Metode EDR	Strength	Weakness	Accuracy
Ebrahim et al. [49]	EDR based on multiresolution discrete wavelet transformation (DWT-MR) and spectral methods	Can estimate the respiration frequency of ECG signals using adaptive signal decomposition techniques and effective spectrum analysis techniques	Requires selection of wavelet bases suitable for ECG signal characteristics, selection of optimal decomposition levels, and accurate R-peak detection process	This is not stated outright but is demonstrated by a graph contrasting the estimated and actual respiration rates
Berset et al. [50]	EDR based on multiresolution discrete wavelet transformation (DWT-MR) on DB4M wavelet base	Can produce smoother and more accurate EDR signals from dynamically varying ECG signals by using wavelet bases that are more in line with ECG signal characteristics	Requires optimal selection of decomposition levels and proper R-peak detection processes	0.98 (mean correlation coefficient between EDR and respiratory signal)
Arunachalam et al. [51]	EDR based on multiresolution discrete wavelet transformation (DWT-MR) on Symlet-4 wavelet base	Can produce smoother and more accurate EDR signals from dynamically varying ECG signals by using wavelet bases that are more in line with ECG signal characteristics	Requires optimal selection of decomposition levels and proper R-peak detection processes	0.97 (mean correlation coefficient between EDR and respiratory signal)

(continued)

Table 3 (continued)

Author	Metode EDR	Strength	Weakness	Accuracy
Imam et al. [52]	EDR based on multiresolution discrete wavelet transformation (DWT-MR) based on the Coiflet-3 wavelet	Can produce smoother and more accurate EDR signals from dynamically varying ECG signals by using wavelet bases that are more in line with ECG signal characteristics	Requires optimal selection of decomposition levels and accurate R-peak detection processes	0.96 (mean correlation coefficient between EDR and respiratory signal)
Charlton et al. [53]	EDR based on BW, AM, FM, RR interval variability, and spectral methods of ECG and PPG	Can extract respiration information from ECG and PPG signals that are easy and routinely measured	Requires accurate detection of R peaks and PPG peaks, optimal parameter selection, and effective assessment of data quality	0.77 (mean correlation coefficient between RR estimated by EDR and measured by IP)

approach has the potential to replace the need for extra sensors for detecting respiration signals from ECG signals. However, the EDR approach still has several obstacles and limits that must be addressed to improve its effectiveness and applicability in clinical practice. These obstacles and restrictions include appropriate parameter selection, effective data quality assessment, validation with abnormal or pathological cardiac recording data, and integration with other analysis techniques such as HRV or PPG.

5 Conclusion

The objective of this study, which is to examine the efficacy of several ECG-derived respiratory methods in reliably measuring respiratory parameters, is commendable. This study examines four decomposition techniques, namely empirical mode decomposition (EMD), variational mode decomposition (VMD), empirical wavelet transform (EWT), and kernel principal component analysis (KPCA). These techniques split the ECG signal into several frequency components, including the breathing component, and then extracted it based on various parameters. The research also compares the reliability of various methods by measuring their correlation and coherence coefficients.

With correlation coefficients of 0.91 and 0.94, respectively, and coherence coefficients of 0.95 and 0.97, respectively, this research suggests that EWT and KPCA perform the best in precision and robustness. This study reveals that EWT is the most effective and efficient decomposition approach for extracting respiration signals from ECG data, as it can adapt to the dynamically varying features of ECG signals, reduce noise, and produce smooth and accurate EDR signals. This study also suggests that decomposition approaches can provide a non-invasive and continuous monitoring approach for respiratory parameters; however, they confront obstacles and limits that must be addressed to increase their clinical applicability.

This paper contributes to the field of medical signal processing by presenting a thorough and critical evaluation of numerous ECG-derived respiratory approaches based on decomposition techniques. This study examines the benefits and drawbacks of each technique, as well as the trade-offs between precision and complexity. This work also gives useful insights into the applicability of these techniques in various healthcare settings, including sleep apnea identification, heart rate variability analysis, and physiological monitoring. This paper also proposes future research topics for enhancing the efficacy and reliability of these methods, including parameter selection, data quality assessment, validation using abnormal or diseased cardiac data, and integration with other analysis techniques.

This work contributes to a comprehensive and enlightening analysis of ECG-derived respiratory signal extraction techniques. For future studies, researchers and health professionals can determine the optimal approach for specific applications by balancing precision and complexity.

References

1. Wehrwein EA, Orer HS, Barman SM (2016) Overview of the anatomy, physiology, and pharmacology of the autonomic nervous system. Regulation 37:125
2. Richter M, Wright RA (2020) Autonomic Nervous System (ANS) BT—encyclopedia of behavioral medicine. https://doi.org/10.1007/978-3-030-39903-0_790
3. Sohrt-Petersen L (2013) Evaluation of algorithms for ECG derived respiration in the context of heart rate variability studies, 1–77
4. Oser MAM (2008) Comparison of respiratory rates derived from heart rate variability, ECG amplitude , and nasal/oral airflow. Ann Biomed Eng 36:2085–2094. https://doi.org/10.1007/s10439-008-9580-2
5. Cikach FS Jr, Dweik RA (2012) Cardiovascular biomarkers in exhaled breath. Prog Cardiovasc Dis 55:34–43
6. Weis JS, Weis JS (2014) Respiration and metabolism. Physiol Dev Behav Eff Mar Pollut, 65–95
7. Ionescu CM (2013) The human respiratory system: an analysis of the interplay between anatomy, structure, breathing and fractal dynamics. Springer Science & Business Media
8. Harada D, Asanoi H, Takagawa J, Ishise H, Ueno H, Oda Y, Goso Y, Joho S, Inoue H (2014) Slow and deep respiration suppresses steady-state sympathetic nerve activity in patients with chronic heart failure: from modeling to clinical application. Am J Physiol Circ Physiol 307:H1159–H1168
9. Molkov YI, Zoccal DB, Baekey DM, Abdala APL, Machado BH, Dick TE, Paton JFR, Rybak IA (2014) Physiological and pathophysiological interactions between the respiratory central pattern generator and the sympathetic nervous system. Prog Brain Res 212:1–23

10. Deepak KK (2002) The role of autonomic nervous system in rapid breathing practices. In: Proceedings: science of breath. international symposium on sudarshan kriya, pranayam and consciousness, All India Institute of Medical Sciences, New Delhi, pp 42–45
11. Pérez-Riera AR, Barbosa-Barros R, Daminello-Raimundo R, de Abreu LC, Nikus K (2019) Current aspects of the basic concepts of the electrophysiology of the sinoatrial node. J Electrocardiol 57:112–118
12. Unudurthi SD, Wolf RM, Hund TJ (2014) Role of sinoatrial node architecture in maintaining a balanced source-sink relationship and synchronous cardiac pacemaking. Front Physiol 5:446
13. Kashou AH, Basit H, Chhabra L (2017) Physiology, sinoatrial node
14. Askenasy JJM, Askenasy N (1996) Inhibition of muscle sympathetic nerve activity during yawning. Clin Auton Res 6:237–239
15. Jindal GD, Sawant MS, Pande JA, Rohini A, Jadhwar P, Naik BB, Deshpande AK (2016) Heart rate variability: objective assessment of autonomic nervous system. MGM J Med Sci 3:198–205
16. Chen C, Dicarlo SE (1998) Endurance exercise training-induced resting Bradycardia: a brief review. Res Sport Med Int J 8:37–77
17. Levy MN, Martin PJ (1989) Autonomic neural control of cardiac function. In: Physiology and pathophysiology of the heart. Springer, pp 361–379
18. McCormick B (2008) Update in Anaesthesia: Editorial, vol 24(3)
19. Moody GB, Mark RG, Zoccola A, Mantero S (1985) Derivation of respiratory signals from multi-lead ECGs. Comput Cardiol 1985(12):113–116. https://doi.org/10.1109/FBIE.2008.41
20. Negi S, Singh RK, Anoop CS (2017) Development of a real-time breathing-rate monitor using difference operation method and adaptive windowing on dry-electrode ECG signal. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, EMBS 2017, pp 1529–1533. https://doi.org/10.1109/EMBC.2017.8037127
21. Mirmohamadsadeghi L, Vesin JM (2015) Estimating the real-time respiratory rate from the ECG with a bank of notch filters. Comput Cardiol 2010(42):581–584. https://doi.org/10.1109/CIC.2015.7410977
22. Maghfiroh AM, Arifin A, Sardjono TA (2019) Wavelet-based respiratory rate estimation using electrocardiogram, 354–359
23. Arunachalam SP, Brown LF (2009) Real-time estimation of the ECG-Derived Respiration (EDR) signal using a new algorithm for baseline wander noise removal. In: Proceedings of the 31st annual international conference of the IEEE engineering in medicine and biology society: engineering the future of biomedicine, EMBC 2009, vol 1, pp 5681–5684. https://doi.org/10.1109/IEMBS.2009.5333113
24. Leif S, Bail R, Laguna P (2006) ECG-derived respiratory frequency estimation. In: Advanced methods and tools for ECG data analysis, pp 215–244
25. Hahn G, Sipinková I, Baisch F, Hellige G (1995) Changes in the thoracic impedance distribution under different ventilatory conditions. Physiol Meas 16:A161–A173. https://doi.org/10.1088/0967-3334/16/3a/016
26. Moody GB, Mark RG, Zocola A, Mantero S (1985) Derivation of respiratory signals from multi-lead ecgs. Comput Cardiol, 113–116
27. Maghfiroh A, Musvika S, Abdullayev V (2022) Performance comparison of ECG bio-amplifier between single and bi-polar supply using spectrum analysis based on fast fourier transform. Indones J Electron Electromed Eng Med Informatics 4. https://doi.org/10.35882/ijeeemi.v4i4.248
28. Maghfiroh AM, Musvika SD, Wakidi LF, Titisari D, Setiawan SY, Amrinsani F, Azhari DH (2022) State-of-the-art method denoising electrocardiogram signal: a review BT. In: Proceedings of the 2nd international conference on electronics, biomedical engineering, and health informatics
29. Langley P, Bowers EJ, Murray A (2009) Principal component analysis as a tool for analyzing beat-to-beat changes in ECG features: application to ECG-derived respiration. IEEE Trans Biomed Eng 57:821–829

30. Janbakhshi P, Shamsollahi MB (2018) Sleep apnea detection from single-lead ECG using features based on ECG-derived respiration (EDR) signals. Irbm 39:206–218
31. Widjaja D, Taelman J, Vandeput S, Braeken MAKA, Otte RA, Van den Bergh BRH, Van Huffel S (2010) ECG-derived respiration: comparison and new measures for respiratory variability. In: 2010 computing in cardiology. IEEE, pp 149–152
32. Zarei A, Asl BM (2020) Automatic classification of apnea and normal subjects using new features extracted from HRV and ECG-derived respiration signals. Biomed Signal Process Control 59:101927
33. Rate EH, Using E, Empirical E, Fusion SD (2021) Estimating heart rate and respiratory rate from a single lead
34. Fan SH, Chou CC, Chen WC, Fang WC (2015) Real-time obstructive sleep apnea detection from frequency analysis of EDR and HRV using Lomb Periodogram. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology, EMBS, 2015-Novem, pp 5989–5992. https://doi.org/10.1109/EMBC.2015.7319756
35. Chen H, Maharatna K (2019) An automatic R-peak detection method based on hierarchical clustering. In: BioCAS 2019–proceedings of the biomedical circuits and systems conference, pp 1–4. https://doi.org/10.1109/BIOCAS.2019.8919208
36. Nazari M, Sakhaei SM (2017) Variational mode extraction: a new efficient method to derive respiratory signals from ECG. IEEE J Biomed Heal Informatics 2194. https://doi.org/10.1109/JBHI.2017.2734074
37. Labate D, Foresta FL, Occhiuto G, Morabito FC, Lay-Ekuakille A, Vergallo P (2013) Empirical mode decomposition vs. wavelet decomposition for the extraction of respiratory signal from single-channel ECG: a comparison. IEEE Sens J 13:2666–2674. https://doi.org/10.1109/JSEN.2013.2257742
38. Campolo M, Labate D, La Foresta F, Morabito FC, Lay-Ekuakille A, Vergallo P (2011) ECG-derived respiratory signal using Empirical Mode Decomposition. In: MeMeA 2011—Proceedings of the 2011 IEEE international symposium on medical measurements and applications, pp 399–403. https://doi.org/10.1109/MeMeA.2011.5966727
39. Nazari M, Sakhaei SM (2015) An efficient method for extracting respiratory activity from single-lead-ECG based on variational mode decomposition, 25–27
40. Sharma H, Sharma KK (2017) Application of iterated Hilbert transform for deriving respiratory signal from single-lead ECG. In: Proceedings of the India international conference on information processing, IICIP 2016. https://doi.org/10.1109/IICIP.2016.7975307
41. Avci C, Delibaşoğlu I, Akbaş A (2012) Sleep apnea detection using wavelet analysis of ECG derived respiratory signal. In: 2012 international conference on biomedical engineering, ICoBE 2012, pp 272–275. https://doi.org/10.1109/ICoBE.2012.6179019
42. Sharma H, Sharma KK, Lata O (2015) Respiratory rate extraction from single-lead ECG using homomorphic fi ltering. Comput Biol Med 59:80–86. https://doi.org/10.1016/j.compbiomed.2015.01.024
43. Schmidt M, Schumann A, Müller J, Bär K-J, Rose G (2017) ECG derived respiration: comparison of time-domain approaches and application to altered breathing patterns of patients with schizophrenia. Physiol Meas 38:601–615. https://doi.org/10.1088/1361-6579/aa5feb
44. Tiinanen S, Noponen K, Tulppo M, Kiviniemi A, Seppänen T (2015) ECG-derived respiration methods: adapted ICA and PCA 37:512–517. https://doi.org/10.1016/j.medengphy.2015.03.004
45. Yang Q, Zou L, Wei K, Liu G (2022) Obstructive sleep apnea detection from single-lead electrocardiogram signals using one-dimensional squeeze-and-excitation residual group network. Comput Biol Med 140:105124. https://doi.org/10.1016/j.compbiomed.2021.105124
46. Tyagi PK, Agrawal D (2023) Biomedical Signal Processing and Control Automatic detection of sleep apnea from single-lead ECG signal using enhanced-deep belief network model. Biomed Signal Process Control 80:104401. https://doi.org/10.1016/j.bspc.2022.104401
47. Wang Z, Fang Y, Zhou D, Li K, Cointet C, Liu H (2020) Ultrasonography and electromyography based hand motion intention recognition for a trans-radial amputee : A case study. Med Eng Phys 75:45–48. https://doi.org/10.1016/j.medengphy.2019.11.005

48. Khandoker AH, Karmakar CK, Palaniswami M (2009) Automated recognition of patients with obstructive sleep apnoea using wavelet-based features of electrocardiogram recordings 39:88–96. https://doi.org/10.1016/j.compbiomed.2008.11.003
49. Ebrahimi F, Setarehdan S, Nazeran H (2015) Biomedical Signal Processing and Control Automatic sleep staging by simultaneous analysis of ECG and respiratory signals in long epochs. Biomed Signal Process Control 18:69–79. https://doi.org/10.1016/j.bspc.2014.12.003
50. Berset T, Romero I, Young A, Penders J (2012) Robust heart rhythm calculation and respiration rate estimation in ambulatory ECG monitoring, vol 25, pp 400–403
51. Arunachalam SP, Brown L (2009) Real-time estimation of the ECG-Derived Respiration (EDR) signal using a new algorithm for baseline wander noise removal, vol 1, pp 5681–5684
52. Imam MH, Karmakar CK, Khandoker AH, Palaniswami M (2013) Effect of using ECG derived respiration (EDR) signal in linear parametric QT-RR modeling. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, EMBS, pp 1968–1971. https://doi.org/10.1109/EMBC.2013.6609914
53. Charlton PH, Bonnici T, Tarassenko L (2016) An assessment of algorithms to estimate respiratory rate from the electrocardiogram and An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram. https://doi.org/10.1088/0967-3334/37/4/610

A Deep CNN-Based Approach for 10-Class with Two-Channel EMG Signal Classification



Triwiyanto, Endro Yulianto, Triana Rahmawati, and Rifai Chai

Abstract Electromyography (EMG) signals have emerged as vital tools for prosthetic control, motor function assessment, and rehabilitation technology. Traditional methods, relying on manual feature extraction, struggle with high-dimensional EMG data. Deep learning, specifically Convolutional Neural Networks (CNNs), has shown promise in addressing these challenges. However, there is a need to explore CNNs' effectiveness, optimal architecture, real-time processing, and data augmentation for EMG signal classification. This study aimed to introduce a CNN-based approach tailored to classify ten-class, two-channel EMG signals. Eight participants provided EMG data by performing ten distinct finger and hand movements. Data were collected using two-channel EMG sensors, amplified, sampled at 4000 Hz, and filtered. A CNN architecture with eight layers was designed, incorporating one-dimensional convolutional layers, max pooling, global average pooling, dropout layers, and a dense output layer. The CNN-based approach demonstrated promising results. It exhibited high accuracy in classifying EMG signals for most classes, with an overall accuracy of approximately 91%. It excelled in the thumb (THU) and thumb-ring (TH-R) classes. However, some classes, like thumb-middle (TH-M) and thumb-index (TH-I), showed room for improvement in terms of recall. This study introduced an innovative CNN-based approach to tackle the challenges of classifying ten-class, two-channel EMG signals. It addressed critical research gaps and significantly advanced EMG signal processing, particularly for prosthetic control and rehabilitation technology.

Keywords Electromyography · Classification · CNN · Deep learning

Triwiyanto (✉) · E. Yulianto · T. Rahmawati
Department of Medical Electronics Technology, Poltekkes Kemenkes Surabaya, Surabaya, Indonesia
e-mail: triwiyanto@ieee.org

R. Chai
School of Science, Computing and Engineering Technologies, Swinburne University of Technology, Melbourne, Australia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

T. Triwiyanto et al. (eds.), *Proceedings of the 4th International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1182, https://doi.org/10.1007/978-981-97-1463-6_46

1 Introduction

The field of biomedical signal processing has undergone a remarkable transformation in recent years, driven by the ever-increasing demand for advanced techniques capable of unraveling the complexities of physiological data. Among these signals, electromyography (EMG) signals have emerged as a crucial element, offering insights into muscle activity that have profound implications for prosthetic control, motor function assessment, and rehabilitation technology [1–4]. Central to these applications is the challenge of accurately classifying EMG signals, a challenge that we aim to address in this study using a deep Convolutional Neural Network (CNN)-based approach. EMG signals are the electrical manifestations of muscle contractions captured through surface electrodes attached to the skin. These signals hold a wealth of information about muscle activity and can be harnessed to infer a user's intended movements, making them indispensable for applications ranging from prosthetic limb control to assisting individuals with motor impairments [5–7].

In the context of prosthetic control, the classification of EMG signals into distinct categories is pivotal. Traditionally, pattern recognition methods have been employed for this purpose. However, these techniques often struggle to achieve high accuracy, especially when confronted with the complexity of scenarios involving numerous classes, as is the case in the 10-class classification problem we address in this study [8–10]. The pursuit of improved EMG signal classification accuracy has led researchers to explore a spectrum of approaches, spanning traditional machine learning techniques and cutting-edge deep learning methods. Traditional methods such as Support Vector Machines (SVMs), k-Nearest Neighbors (k-NN), and Random Forests have been employed for EMG signal classification. However, these approaches typically rely on manual feature extraction, which can limit their effectiveness when dealing with high-dimensional and intricate EMG data [11–13]. Deep learning has revolutionized the field in recent years by offering promising solutions to these challenges. Initially designed for image analysis, convolutional Neural Networks (CNNs) have demonstrated remarkable success in various signal-processing tasks, including speech recognition, natural language processing, and medical image analysis. Importantly, CNNs have shown great potential in EMG signal classification, offering the prospect of improved accuracy and robustness [14–17].

While significant strides have been made in EMG signal classification, several critical challenges and research gaps persist: (a) EMG signals inherently possess high dimensionality, capturing intricate patterns of muscle activity. Extracting these patterns accurately remains a significant challenge. We aim to explore how CNNs can effectively extract discriminative features from this complex data, (b) Real-world applications often require the classification of EMG signals into multiple classes, a scenario that frequently exceeds binary or few-class scenarios. Our study focuses on the specific challenges associated with 10-class classification, a context demanding novel solutions, (c) The design of CNN architectures tailored to EMG signal classification is critical for achieving high accuracy. Determining the optimal number

of layers, filter sizes, and pooling strategies is an ongoing research question that we seek to address, and (d) Many applications, particularly in prosthetic control, demand real-time processing of EMG signals. Developing computationally efficient CNN models capable of real-time operation without sacrificing accuracy is a formidable research challenge that we aim to tackle. Adequate training data for EMG signal classification can be scarce, especially for specific classes. Innovating effective data augmentation techniques to mitigate this limitation is an important research direction that we will explore. The field of EMG signal classification stands at a critical juncture where the integration of deep learning, particularly CNNs, holds immense potential to enhance the accuracy and applicability of EMG-based assistive technologies. This paper embarks on a transformative journey, introducing a novel approach that harnesses the power of deep CNNs to address the multifaceted challenges associated with 10-class classification of two-channel EMG signals. By doing so, we aim to make a meaningful contribution to the domains of human–computer interaction and assistive technology [18–20]. The three most significant contributions of this study are:

1. This study introduces a deep Convolutional Neural Network (CNN)-based approach specifically tailored to classify two-channel EMG signals into ten distinct classes. The utilization of CNNs for EMG signal classification is a notable contribution, as it harnesses the power of deep learning to improve accuracy and robustness in a complex multiclass scenario.
2. The study actively identifies and addresses crucial research gaps in the field of EMG signal classification. It takes on the challenges of handling high-dimensional data, optimizing CNN architectures, enabling real-time applications, and innovating data augmentation techniques. By tackling these gaps, the research significantly advances the state-of-the-art in EMG signal processing.
3. The proposed deep CNN-based approach has practical implications for assistive technologies, including prosthetic control and rehabilitation. By improving the accuracy of EMG signal classification, this research can enhance the user experience and functionality of assistive devices, ultimately contributing to the betterment of individuals with motor impairments and those in need of prosthetic solutions. This advancement in human–computer interaction is a substantial contribution to the field.

2 Method

2.1 Dataset

Eight participants between the ages of 20 and 35 were recruited to carry out the required finger movements. There were six men and two women in the group. Each individual had a normal set of limbs and showed no signs of neurological or muscular abnormalities. All participants gave their informed consent before beginning the trial.

The participants were seated in an armchair, their arms immobilized in a stable position and supported securely. This was done to reduce the impact of different limb positions on the electromyography (EMG) data generated (Scheme, Founger, Stavdahl, Chan, & Englehart, 2010). Two EMG channels, especially the Delsys DE 2.x series EMG sensors, were used to collect the EMG data. Delsys Inc.'s Bagnoli Desk-top EMG Systems were used to process this data. A 2-slot adhesive skin interface was used to secure each sensor's attachment to the skin.

2.2 Data Collection

Following collecting EMG signals from these carefully positioned electrodes, a Delsys Bagnoli-8 amplifier was utilized to amplify the data. This amplifier was chosen due to its reliability and accuracy, and it can achieve a total gain of 1000 by multiplying the signals. Because it improves the signals, this step is essential because it prepares them for subsequent processing and analysis. After the amplification was completed, a 12-bit analog-to-digital converter, more precisely the National Instruments BNC-2090, was utilized to sample the signal at a rate of 4000 Hz. The conversion procedure was designed to provide the maximum possible level of accuracy, hence selecting this particular converter and sample rate was necessary.

After that, the signal data were gathered with the help of the Delsys EMGWorks Acquisition program, which is a reliable instrument in the field due to the robustness and user-friendliness of its interface. The EMG signals were filtered as the following stage in the process. They were subjected to bandpass filtering between 20 and 450 Hz, and a notch filter was also used to eliminate any potential interference from the 50 Hz line. The signal must first go through this necessary filtering procedure, which removes any noise or frequencies that aren't needed while retaining just the data that's actually useful for analysis. The experiment was designed to examine 10 different categories of solo and combined finger motions. These included flexion of each individual finger, including the thumb (T), index (I), middle (M), ring (R), and little finger (L), as well as combined movements such as flexion of the thumb and index, flexion of the thumb and middle, flexion of the thumb and ring, and flexion of the thumb and little finger. The final action was to shut one's hand (HC).

These motions were selected because they represent a broad variety of finger and hand movements often utilized in day-to-day tasks. Each movement was meticulously tracked and documented, and the resulting dataset contains several precise numbers for additional research purposes. A document was also supplied for extra details, to provide an in-depth description of each movement and its related EMG signal. This experiment makes a substantial contribution to the field of EMG signal analysis, which is the subject of the study. The painstaking process of electrode installation, signal amplification, conversion, collection, and filtering ensures the accuracy and dependability of the data acquired. The emphasis on individual and combined finger motions gives vital insights into how these movements might be recognized and evaluated using EMG data. The ramifications of this research are quite broad and

may be found in a variety of sectors including prostheses, rehabilitation, and human–computer interaction.

2.3 *Proposed Method*

In this study, the proposed method to complete the 10-class classification process is to use two channels of EMG signals, as described in Fig. 1. The proposed method to solve this problem is to use the convolution neural network (CNN) architecture. This classification is composed of two-channel EMG signal inputs and CNNs. Furthermore, the CNN classifier consists of 8 layers, namely L1, L2, L3, L4, L5, L6, L7 and L8. The L1, L2, L4, and L5 layers are one-dimensional convolution neural network layers with 200 filters randomly configured initially.

The kernel used in the convolution process has dimensions of 4×1, which is the convolution between the kernel filter and the EMG signal done sequentially. Layer 3 is Max pooling, which aims to reduce the number of convolution outputs by choosing the maximum convolution result. Max pooling output at layer L3 produces an output of 98 nodes. Global avarage pooling (GAP) reproduces each filter into one

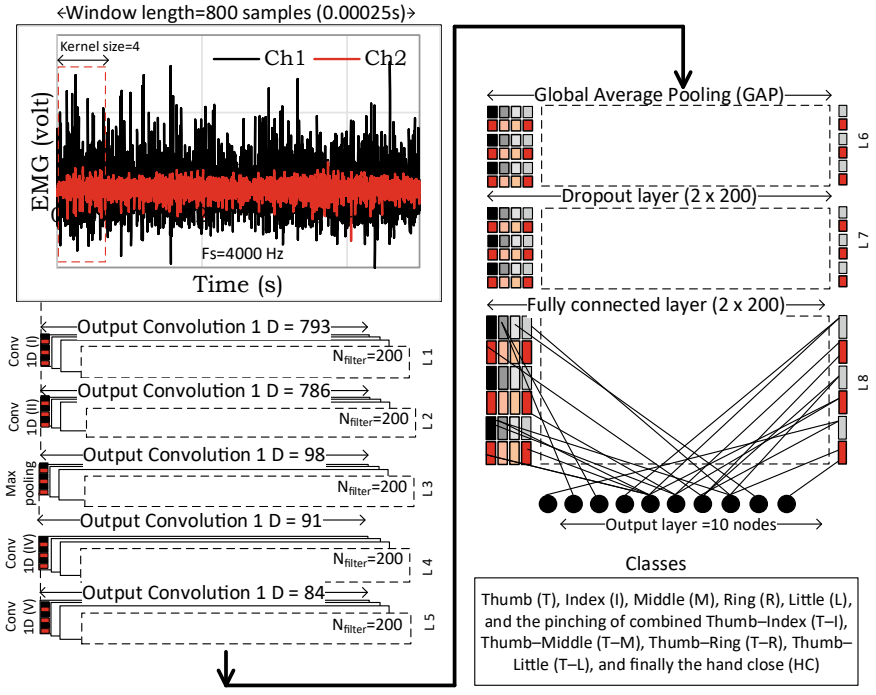


Fig. 1 CNN model architecture

Table 1 CNN model structure

Layer (type)	Output Shape	Param #
reshape (Reshape)	(None, 800, 2)	0
conv1d (Conv1D)	(None, 793, 200)	3400
conv1d_1 (Conv1D)	(None, 786, 200)	320,200
max_pooling1d (MaxPooling1D)	(None, 98, 200)	0
conv1d_2 (Conv1D)	(None, 91, 200)	320,200
conv1d_3 (Conv1D)	(None, 84, 200)	320,200
global_average_pooling1d (Gl)	(None, 200)	0
dropout (Dropout)	(None, 200)	0
dense (Dense)	(None, 10)	2010

node. At the end of layer L8, all nodes are connected to the output node (10 nodes). Output nodes are in accordance with the number of classes to be used, which is as many as ten classes (Thumb (T), Index (I), Middle (M), Ring (R), Little (L), and the pinching of combined Thumb–Index (T–I), Thumb–Middle (T–M), Thumb–Ring (T–R), Thumb–Little (T–L), and finally the hand close (HC)).

The given model (Table 1) describes a Convolutional Neural Network (CNN) designed for 1-dimensional input data. The network begins with a reshape layer that adjusts the input data into a 3D tensor format, suitable for processing by subsequent layers. This is followed by convolutional layers, the heart of the network, responsible for learning local patterns in the input data. The first Conv1D layer transforms the tensor using 3,400 parameters, while the second and third Conv1D layers further process the data using 320,200 parameters each. A MaxPooling1D layer is then applied to reduce the spatial dimensions of the output from the preceding convolutional layer. This abstraction helps to reduce overfitting and computational cost. The network also includes a GlobalAveragePooling1D layer that computes the average output of each feature map in the previous layer, effectively reducing the total number of parameters in the model and improving model generalization. To prevent overfitting, a dropout layer is included which randomly sets a fraction of input units to 0 during each update in training time. Finally, a dense layer is used to output predictions for each class. This fully connected layer connects every neuron to every neuron in the previous layer and uses 2,010 parameters.

3 Result

The EMG signal is intercepted using a two-channel electrode on the radial carpi flexor (CH1) and ulnar carpi extensor (CH2). EMG signals (Fig. 2) show different signal patterns for each movement (Thumb (T), Index (I), Middle (M), Ring (R), Little (L), and the pinching of combined Thumb–Index (T–I), Thumb–Middle (T–M), Thumb–Ring (T–R), Thumb–Little (T–L), and finally the hand close (HC)). Figure 2a

shows that during hand close (HC) movement, both EMG signals exhibit the same bio-electrical activity as shown by an increase in the signal clumping amplitude. Different circumstances occur in Fig. 2b, namely in the movement of index (I), little (L), thumb-index (T-I), thumb-little (T-L), and thumb (T-T). EMG signal activity produces different amplitudes where CH1 has higher activity compared to activity at CH2.

Figure 3 is a line graph that gives a visual representation of a machine learning model's performance over 50 training epochs. The training epochs are represented along the x-axis, while the accuracy and loss values are shown along the y-axis and range from 0 to 1 respectively. On the graph, you'll notice that there are four lines. Each of these lines represents a different measure. The solid green line indicates the loss on the training data, whereas the dashed green line represents the model's accuracy based on the training data. Similarly, the red line with dashes shows the accuracy on the validation data, but the red line without dashes represents the loss on the validation data. According to the graph, both accuracy lines (one dashed green and one dashed red) exhibit a decreasing trend over time, which indicates that the accuracy of the model is improving with each epoch.

On the other hand, both loss lines (solid green and red) exhibit an upward trend, indicating that the model's loss decreases with each epoch. This is shown by the fact that the trend is upward. This graph offers a clear and concise visual picture of how the model's performance develops over time in terms of accuracy and loss as it is being trained. It demonstrates that as more epochs are used to train the model, it gets more accurate and lowers its loss.

The performance of a classification model on a set of data for which the real values are already known can be described with the use of a table known as a confusion matrix (Fig. 4). The matrix in question is a 10 by 10 grid, which indicates that the model was charged with classifying cases into one of nine potential classes. There are also nine alternative classes. CLO, IND, LIT, MID, RIN, THI, THL, THM, THR, and THU are the labels that are assigned to these courses. Each cell in the matrix represents the number of predictions the model has produced for each possible combination of true and predicted labels. The color scale runs from blue all the way up to red, with blue representing a low number of forecasts and red representing a large number of predictions. The diagonally aligned cells show correct predictions, as the anticipated class corresponds to the actual class in those situations. For example, there were 1971 occurrences in which "THM" served as both the true and anticipated classes. This is the greatest number of accurate predictions that the model has ever produced. The cells located outside the diagonal show situations in which the predicted class does not match the actual class, often known as inaccurate predictions. For example, there was not a single occurrence in which the correct class was "CLO" while the anticipated class was "IND." This leads one to believe that the model has never wrongly categorized a case of "CLO" as an instance of "IND." A confusion matrix is a detailed examination of a model's performance that shows us how many predictions were accurate (represented by the diagonal cells) and the different types of errors that the model made (represented by the off-diagonal cells).

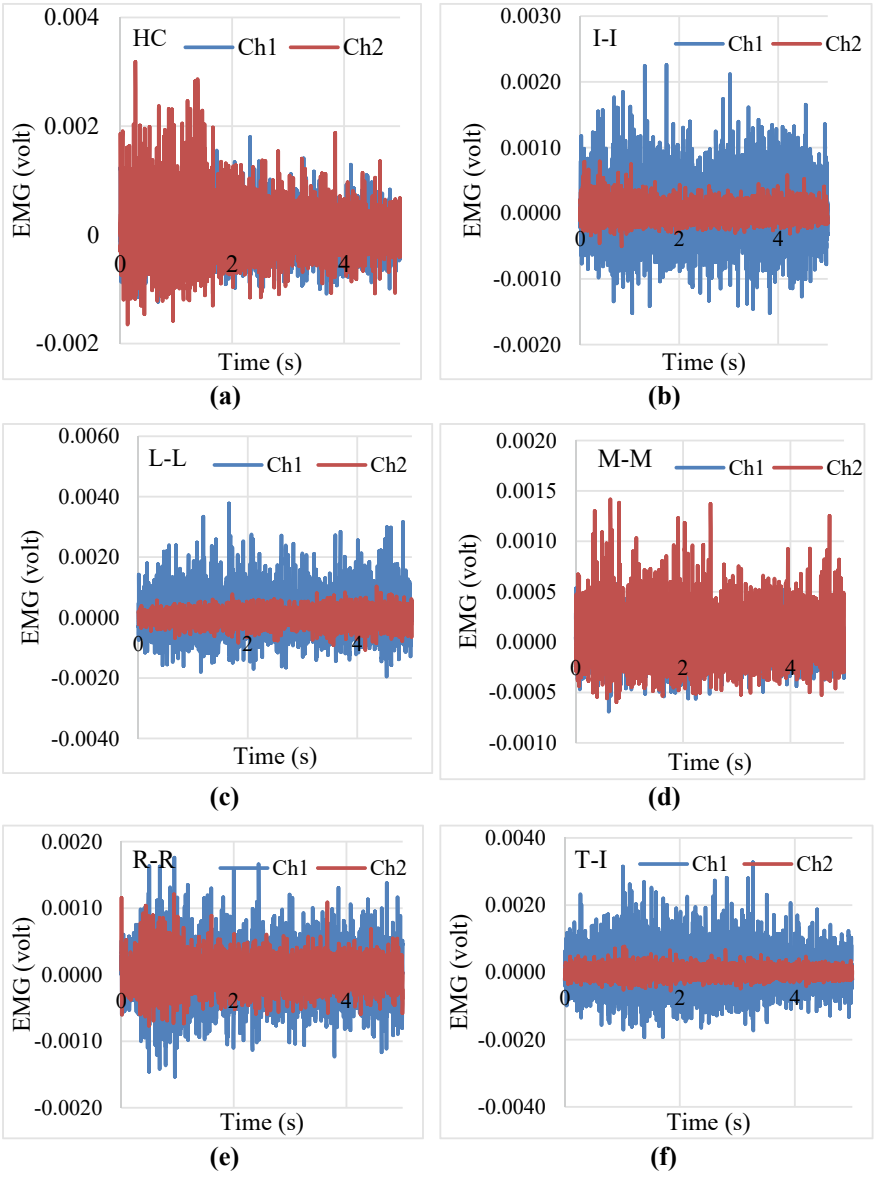


Fig. 2 Two channel EMG signal for ten motion: **a** hand close (HC), **b** Index (I), **c** Little (L), **d** Middle (M), **e** Ring (R), **f** Thumb–Index (T–I), **g** Thumb–Little (T–L), **h** Thumb–Middle (T–M), **i** Thumb–Ring (T–R), **j** (Thumb (T)

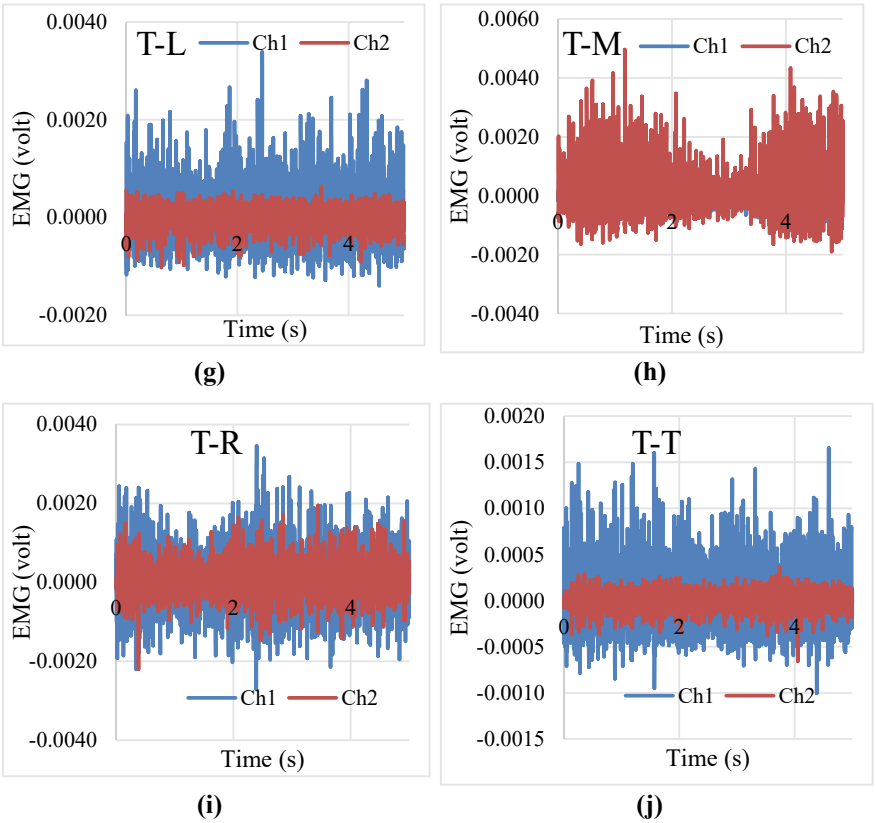


Fig. 2 (continued)

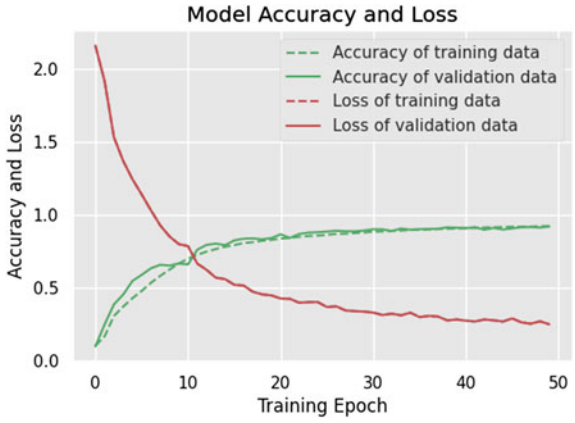


Fig. 3 Model accuracy and loss

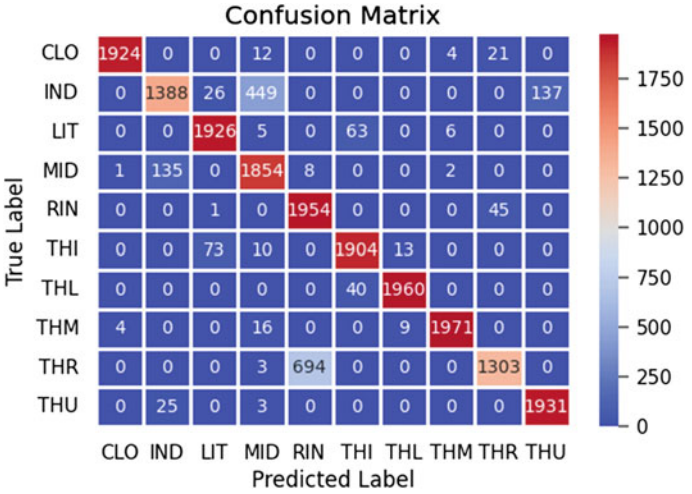


Fig. 4 Confusion matrix for test data

The classification report provided comprehensively evaluates a classification model's performance on the test data (Table 2). The model exhibits perfect precision for class 0, implying that every instance it predicted as class 0 was indeed class 0. It identified 98% of all actual instances of class 0, as indicated by the recall of 0.98. The F1-score, the harmonic mean of precision and recall, is 0.99, demonstrating excellent performance for this class. For class 1, the model's precision is 0.9, indicating that 90% of the instances it predicted as class 1 were correct. However, the recall is lower at 0.69, suggesting that it only identified 69% of all actual instances of class 1. This results in an F1-score of 0.78. The model shows high precision and recall for class 2 (0.95 and 0.96 respectively), leading to an F1-score of 0.96. Despite a lower precision of 0.79 for class 3, the model has a high recall of 0.93, resulting in an F1-score of 0.85. Class 4 has a lower precision of 0.74 but a very high recall of 0.98, leading to an F1-score of 0.84. Classes 5,6,7 exhibit high precision, recall and F1-scores, indicating excellent performance. Class 8 has a high precision but lower recall (0.65), resulting in an F1-score of 0.77. Finally, class 9 shows high precision and very high recall (0.99), leading to an F1-score of 0.96. The model's overall accuracy on the test data is approximately 91%. The macro-average and weighted average for precision, recall and F1-score are also around this value (92% and 91% respectively). These averages consider all classes equally (macro) or based on their support i.e., the number of instances in each class (weighted). In conclusion, while the model performs well on most classes, there are opportunities for improvement in classes with lower recall.

Table 2 Classification report for test data

Class	Precision	Recall	F1-	Score support
0	1	0.98	0.99	1961
1	0.9	0.69	0.78	2000
2	0.95	0.96	0.96	2000
3	0.79	0.93	0.85	2000
4	0.74	0.98	0.84	2000
5	0.95	0.95	0.95	2000
6	0.99	0.98	0.98	2000
7	0.99	0.99	0.99	2000
8	0.95	0.65	0.77	2000
9	0.93	0.99	0.96	1959
Accuracy			0.91	19,920
Macro avg	0.92	0.91	0.91	19,920
Weighted avg	0.92	0.91	0.91	19,920

4 Discussion

The findings of this study offer a thorough understanding of the strengths and weaknesses of a deep convolutional neural network (CNN)-based method for categorizing two-channel electromyography (EMG) signals into 10 different groups. The study's main finding is that performance varies significantly between classes. The model achieved high precision, recall, and F1-scores when categorizing classes like the thumb (T), index (I), middle (M), and ring (R). This shows that the CNN was successful in capturing the distinctive muscle activity patterns linked to these particular finger movements. The exact finger movement detection required for applications like prosthesis control, where these classes may be distinguished with high precision, has important ramifications. On the other hand, other classes had more difficult categorization jobs. Class 1, which represents particular finger movements, displayed a lower recall value, indicating that the model failed to recognize all instances of this class when classifying them. This finding emphasizes how challenging it is to effectively differentiate between these classes' slight variations in muscle activity patterns. It could be important to do additional study and tweak the model in order to increase classification precision for such complex motions. Using average, the CNN-based method on the test data had an accuracy rate of almost 91 percent. This result clearly indicates how well the model performs in differentiating between the ten classes. Although this accuracy is great, it is important to keep in mind that accuracy may not accurately depict all of the model's capabilities, particularly in datasets with imbalances where certain classes have more instances than others. This study follows the general trend in the field of EMG signal processing, where deep learning approaches, particularly CNNs, have gradually superseded classic machine learning strategies like Support Vector Machines

(SVMs), k-Nearest Neighbors (k-NN), and Random Forests. When working with high-dimensional EMG data, traditional methods frequently call for manual feature extraction, which can be constricting. Conversely, this study shows that CNNs can automatically extract discriminative features from complicated EMG data, negating the necessity for feature engineering. This move toward deep learning could considerably improve the accuracy and robustness of categorization. The model's accuracy in classifying EMG signals into various groups, even in a difficult 10-class situation, is an example of the power of deep learning in this field. The study results have significant ramifications for practical applications, particularly in prosthetic control, rehabilitation, and human–computer interface. Assistive devices' usefulness and user experience can be improved with increased EMG signal categorization accuracy. More intuitive and sensitive prosthetic devices can help people with motor disabilities and improve their daily lives.

Several research have investigated the classification of EMG signals using a variety of methodologies, including deep learning and conventional machine learning. Traditional machine learning methods like Support Vector Machines (SVMs) and k-Nearest Neighbors (k-NN) have been used in previous investigations, such those in [11] and [21], to classify EMG signals. These approaches frequently depend on manual feature extraction, which has its drawbacks when working with high-dimensional EMG data. In contrast, the labor-intensive process of feature engineering is avoided in our study by using deep convolutional neural networks (CNNs) to automatically extract discriminative features from complicated EMG data [22]. This move in the direction of deep learning fits with the general trend in the industry and has the potential to improve accuracy and robustness [23] greatly. The high dimensionality of EMG signals must be addressed, which is a major problem. To classify EMG signals, CNNs and other deep learning approaches have been investigated in studies like [10] and [16]. Their common objective is the feature extraction from challenging EMG data. In our study, we contribute to this field by emphasizing the efficient extraction of discriminative features from high-dimensional data. In order to demonstrate the efficacy of the CNN-based strategy in addressing this challenging circumstance, our study examines the specific difficulties connected with 10-class categorization. This adds to the growing body of knowledge about the classification of multiple classes of EMG signals. Current research focuses on how to create CNN architectures that are optimized for classifying EMG signals. Earlier research has investigated several CNN architectures and settings to enhance model performance, including [23] and [24]. The CNN classifier used in our investigation has 8 layers, each with a unique configuration [11]. In order to determine the ideal layer counts, filter sizes, and pooling techniques, we assess how well these layers perform in solving the 10-class classification issue. In studies like [21] and [25], the need for real-time processing of EMG signals is addressed, notably in applications like prosthesis control. The task of increasing computational efficiency without sacrificing accuracy is difficult. Our study recognizes the value of real-time processing and offers findings that open the door to creating CNN models that can operate in real-time without sacrificing accuracy due to their computational efficiency. While

the work offers a starting point, more investigation in this area can improve the stability of EMG signal classification algorithms.

The limitations and weaknesses that should be acknowledged in this study are as follow. The research covered in the previous sections shows important advancements in the categorization of EMG signals. The relatively small and homogenous sample size of participants is one major drawback which should be addressed as is true of all studies. Only eight participants, all between the ages of 20 and 35 who had no known neurological or muscle anomalies were selected for the study. Participants' lack of demographic variety may have hampered the findings' generalizability. A more thorough understanding of the efficacy of the suggested CNN-based strategy for diverse user groups would result from broader participant participation, which would include people of a larger age range and those with varying physical conditions. Participants were immobilised in armchairs for the controlled and optimal data gathering environment. While this controlled environment is excellent for establishing baseline accuracy, it might not accurately reflect the noise and unpredictability found in real-world circumstances. Users come across a variety of circumstances and movements in practical applications like prosthesis control. The robustness and application of the model could be evaluated more accurately by expanding the study to incorporate data gathered in more realistic and dynamic contexts. The study recognizes the value of data augmentation, especially for classes with scant training data. However, it doesn't go into great detail about cutting-edge data augmentation methods. A potential shortcoming is a lack of a thorough investigation of augmentation tactics that could improve the model's performance. A thorough analysis of data augmentation strategies designed for EMG signal categorization could be useful for future study. The study used specialized EMG sensors (EMG sensors from the Delsys DE 2.x series) and an exclusive amplifier (Delsys Bagnoli-8) for data collection and processing. Despite the fact that this configuration was chosen for its dependability and accuracy, the model's generalizability to various EMG sensors and amplifiers has not been addressed. When used in different situations, variations in sensor kinds and attributes may impact a model's performance. The model's capacity to accommodate different EMG sensor setups may be the subject of further study. When analyzing the study's findings and preparing for further research projects on the classification of EMG signals, it is important to keep these limits and flaws in mind. The potential and usefulness of the suggested CNN-based technique can be more fully and thoroughly understood by addressing these constraints.

5 Conclusion

A deep Convolutional Neural Network (CNN)-based methodology was used in this study to improve the classification of two-channel EMG signals into ten separate classifications. This study's main goal was to address major problems with EMG signal classification, such as high dimensionality, multi-class scenarios, the best CNN architecture design, real-time processing, and data augmentation. These study's

results have important and comprehensive findings with average accuracy on test mode 91.0%. First off, the suggested CNN-based approach proved to be remarkably accurate and resilient at classifying EMG signals, especially in the difficult setting of ten-class classification. The model showed how it may enhance precision and dependability in practical applications, such as prosthesis control and rehabilitation technologies. A further benefit of the study was the active identification and filling of significant research gaps, which helped develop EMG signal processing. These results should be the foundation for further research in this area. To guarantee that the model is applicable to other user groups, researchers can investigate further demographic diversification of participant populations. To evaluate the model's performance under realistic circumstances, real-world data collecting in dynamic settings should be given priority. Additionally, thorough research into cutting-edge data augmentation methods designed for EMG signal classification could improve the model's performance. The CNN-based approach's flexibility in handling different EMG sensor kinds and configurations is a last point that deserves investigation.

References

1. Mahdavi FA, Ahmad SA, Marhaban MH, Akhbarzadeh-T M-R (2012) Surface electromyography feature extraction based on wavelet transform. Int J Integr Eng 4:1–7. https://doi.org/10.1109/ICIEA.2013.6566603
2. Triwiyanto T, Wahyunggoro O, Nugroho HA (2018) Time delay neural network to estimate the elbow joint angle based on electromyography. In: 10th international conference on biomedical engineering. https://doi.org/10.1061/40971(310)18
3. Elamvazuthi I, Ling GA, Nurhanim KARK, Vasant P, Parasuraman S (2013) Surface electromyography (sEMG) feature extraction based on Daubechies wavelets. In: 2013 IEEE 8th conference on industrial electronics and applications (ICIEA). IEEE, pp 1492–1495. https://doi.org/10.1109/ICIEA.2013.6566603
4. Nazarpour K, Krasoulis A, Hahne JM (2021) Control of prosthetic hands. In: Control of prosthetic hands: challenges and emerging avenues. Institution of Engineering and Technology, pp 1–13. https://doi.org/10.1049/PBHE022E_ch1
5. Sudarsan LS, Sekaran EC (2012) Design and development of EMG controlled prosthetics limb. Procedia Eng 38:3547–3551. https://doi.org/10.1016/j.proeng.2012.06.409
6. Yadav D, Veer K (2023) Recent trends and challenges of surface electromyography in prosthetic applications. Biomed Eng Lett. https://doi.org/10.1007/S13534-023-00281-Z
7. Triwiyanto T, Rahmawati T, Pudji A, Mak'ruf MR (2022) Syaifudin: deep learning approach in hand motion recognition using electromyography signal: a review. In: Lecture notes in electrical engineering, pp 135–146. https://doi.org/10.1007/978-981-19-1804-9_11
8. Hiraiwa A, Uchida N, Shimohara K (1993) EMG pattern recognition by neural networks for prosthetic fingers control. Artif Intell Real-Time Control 1992:73–79. https://doi.org/10.1016/B978-0-08-041898-8.50018-9
9. Unanyan NN, Belov AA (2021) Low-price prosthetic hand controlled by EMG signals. IFAC-PapersOnLine 54:299–304. https://doi.org/10.1016/j.ifacol.2021.10.463
10. Hiraiwa A, Uchida N, Shimohara K (1992) EMG pattern recognition by neural networks for prosthetic fingers control. Annu Rev Autom Program 17:73–79. https://doi.org/10.1016/s0066-4138(09)91014-x
11. Triwiyanto T, Rahmawati T, Pawana IPA (2019) Feature and muscle selection for an effective hand motion classifier based on electromyography. Indonesian J Electr Eng Inf 7. https://doi.org/10.11591/ijeei.v7i2.867

12. Rechy-ramirez EJ, Hu H (2015) Bio-signal based control in assistive robots : a survey. Digit Commun Netw 1:85–101. https://doi.org/10.1016/j.dcan.2015.02.004
13. Zhang X, Liu Y, Zhang F, Ren J, Sun YL, Yang Q, Huang H (2012) On design and implementation of neural-machine interface for artificial legs. IEEE Trans Ind Inf 8:418–429. https://doi.org/10.1109/TII.2011.2166770
14. Karnam NK, Dubey SR, Turlapaty AC, Gokaraju B (2022) EMGHandNet: A hybrid CNN and Bi-LSTM architecture for hand activity classification using surface EMG signals. Biocybern Biomed Eng 42:325–340. https://doi.org/10.1016/j.bbe.2022.02.005
15. Wang L, Fu J, Chen H, Zheng B (2023) Hand gesture recognition using smooth wavelet packet transformation and hybrid CNN based on surface EMG and accelerometer signal. Biomed Signal Process Control 86. https://doi.org/10.1016/j.bspc.2023.105141
16. Ozdemir MA, Kisa DH, Guren O, Akan A (2022) Hand gesture classification using time–frequency images and transfer learning based on CNN. Biomed Signal Process Control 77. https://doi.org/10.1016/j.bspc.2022.103787
17. Fu YL, Liang KC, Song W, Huang J (2022) A hybrid approach to product prototype usability testing based on surface EMG images and convolutional neural network classification. Comput Methods Programs Biomed 221. https://doi.org/10.1016/j.cmpb.2022.106870
18. Balbinot A, Júnior A, Favieiro G (2013) Decoding arm movements by myoelectric signal and artificial neural networks. https://doi.org/10.4236/ica.2013.41012
19. Raurale SA (2014) Acquisition and processing real-time EMG signals for prosthesis active hand movements. In: 2014 international conference on green computing communication and electrical engineering (ICGCCEE). IEEE, pp 1–6. https://doi.org/10.1109/ICGCCEE.2014.6922225
20. Chen W, Li G, Li N, Wang W, Yu P, Wang R, Xue X, Zhao X, Liu L (2023) Restoring voluntary bimanual activities of patients with chronic hemiparesis through a foot-controlled hand/forearm exoskeleton. IEEE Trans Neural Syst Rehabil Eng 31:769–778. https://doi.org/10.1109/TNSRE.2022.3233631
21. Ruiz-Olaya AF, Quinayas Burgos CA, Londono LT (2019) A low-cost arm robotic platform based on myoelectric control for rehabilitation engineering. In: Chakrabarti S, Saha HN (eds) 2019 IEEE 10th annual ubiquitous computing, electronics and mobile communication conference, UEMCON 2019. Institute of Electrical and Electronics Engineers Inc., pp 929–933. https://doi.org/10.1109/UEMCON47517.2019.8993080
22. Jafarzadeh M, Tadesse Y (2019) Convolutional neural networks for speech controlled prosthetic hands. In: Proceedings of the 2019 1st International Conference on Transdisciplinary AI, TransAI 2019, pp 35–42. https://doi.org/10.1109/TransAI46475.2019.00014
23. Triwiyanto T, Lusiana, Wakidi LF, Amrinsani F (2022) Deep learning approach to detect the Covid-19 infection using chest X-ray image: a review. In: Lecture notes in electrical engineering, pp 237–251. https://doi.org/10.1007/978-981-19-1804-9_19
24. Kiranyaz S, Avci O, Abdeljaber O, Ince T, Gabbouj M, Inman DJ (2021) 1D convolutional neural networks and applications: a survey. Mech Syst Signal Process 151:107398. https://doi.org/10.1016/j.ymssp.2020.107398
25. Roy SS, Samanta K, Chatterjee S, Dey S, Nandi A, Bhowmik R, Mondal S (2020) Hand movement recognition using cross spectrum image analysis of EMG signals: a deep learning approach. In: 2020 national conference on emerging trends on sustainable technology and engineering applications (NCETSTEA), pp 1–5. https://doi.org/10.1109/NCETSTEA48365.2020.9119911