Lecture Notes in Electrical Engineering 1611

Triwiyanto
Achmad Rizal
Wahyu Caesarendra *Editors*

Proceedings of the 5th International Conference on Electronics, Biomedical Engineering, and Health Informatics

ICEBEHI 2024, October 9–10, Surabaya, Indonesia

Springer

Lecture Notes in Electrical Engineering 1611

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, Munich, Germany*
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, 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, 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, USA*
Yong Li, *College of Electrical and Information Engineering, Hunan University, Changsha, China*
Qilian Liang, *Department of Electrical Engineering, University of Texas at Arlington, Arlington, USA*
Ferran Martín, *Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, 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, 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, Italy*
Bijaya Ketan Panigrahi, *Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, India*
Federica Pascucci, *Department di Ingegneria, Università degli Studi Roma Tre, Rome, 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, Beijing, China*
Walter Zamboni, *Department of Computer Engineering, Electrical Engineering and Applied Mathematics, DIEM—Università degli studi di Salerno, Fisciano, Italy*
Kay Chen Tan, *Department of Computing, Hong Kong Polytechnic University, Hong Kong, 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 · Achmad Rizal · Wahyu Caesarendra
Editors

Proceedings of the 5th International Conference on Electronics, Biomedical Engineering, and Health Informatics

ICEBEHI 2024, October 9–10, Surabaya, Indonesia

Springer

Editors
Triwiyanto
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 978-981-95-9232-6 ISBN 978-981-95-9233-3 (eBook)
https://doi.org/10.1007/978-981-95-9233-3

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Singapore Pte Ltd. 2026

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

If disposing of this product, please recycle the paper.

Preface

5th ICEBEHI-2024, the 5th International Conference on Electronics, Biomedical Engineering, and Health Informatics, take place during October 9–10, 2024 on virtual platforms (Zoom app). The conference was organized by Department of Medical Electronics Technology, Health Polytechnic Ministry of Health Surabaya (Poltekkes Kemenkes Surabaya), Surabaya, Indonesia and Co-organized by University College TATI, Malaysia. with aims of bring together as family all leading scientists, academicians, educationist, young scientist, 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-2024 conference was held online through Zoom application. More than 100 participants (presenter and non-presenter) attended the conference, they were from Vietnam, Malaysia, Columbia, Brunei Darussalam, Azerbaijan, South Korea, 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 area of Biomedical Engineering. Each keynote speech lasted 50 min.

ICEBEHI 2024 Conference collects the latest research results and applications on electronics, biomedical engineering, and health informatics. It includes a selection of 38 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 quality of their presentations and communication with the other participants that really made this web conference fruitful. Last but not the 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 — Triwiyanto

Organization

Organizing Committee

Endro Yulianto (Chairman)	Poltekkes Kemenkes Surabaya
Triwiyanto (Technical Program Committee Chair)	Poltekkes Kemenkes Surabaya
Sari Luthfiyah (Secretary I)	Poltekkes Kemenke Surabaya
Anita Miftahul Maghfiroh (Secretary II)	Poltekkes Kemenkes Surabaya
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

Triwiyanto	Poltekkes Kemenkes Surabaya, Indonesia
Achmad Rizal	Telkom University, Bandung, Indonesia
Wahyu Caesarendra	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, UK
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

Contents

Optimized Automated Standing Frame for Postural Therapy in Children with Cerebral Palsy 1
Johanes Wawan Joharwan, Ruzlaini Ghoni, Emanuel Budi Raharjo, Husein Arif Susanto, Ichsan Ramadhan, Damhuji Rifai, and Wasis Nugroho

How the Entrepreneurial Ecosystem Drives Success in Telemedicine Healthcare Businesses: A Systematic Review 12
Puan Ayu Maharani, Mardhatillah Shanti, and Agung Purnomo

Leveraging Medical Informatics for Data-Driven Entrepreneurial Ventures in the Healthcare Sector: A Systematic Review 27
Wafiq Azizah, Riefky Prabowo, and Agung Purnomo

Research Trends in Telemedicine Entrepreneurship: Emerging Mapping from a Global Perspective 40
Noor Anisa Nabila and Agung Purnomo

Biomechanical Simulation and Experimental Study for Total Hip Arthroplasty Placement Tools: Femoral Broach and Acetabular Reamer 53
Dita A. Mayasari, Nonik S. Ayati, Nisa R. Amitasari, Talitha Asmaria, Menik Dwi Kurniatie, Joko Triwardono, and Muhammad S. Utomo

Internet of Things and Entrepreneurship: A Bibliometric Insight into Technological Integration and Business Model Innovation 71
Fira Angelia Gafur and Priska Arindya Purnama

Classification of Basic Taekwondo Kicks Using Frequency Domain Analysis and Machine Learning Algorithms 81
Qoriina Dwi Amalia, Achmad Rizal, Bayu Erfianto, and Istiqomah

Evaluating Ensemble Learning and Neural Network Models Towards Covid-19 Detection Through Cough Recordings 92
William Axel Cuangdinata, Kent Alber Fredson, Alfi Yusrotis Zakiyyah, and Meiliana

Model Development and Features Selection of Brute Force Attack Classification on Application Layer Network 109
Muhammad Afrizal Amrustian, Bita Parga Zen, and Muhammad Ikhwani Saputra

Enhancing COVID-19 Prognostic Accuracy with Machine Learning and Clinical Data Visualization 121
Phuoc-Hai Huynh

Initial Movement of Bionic Foot: Control Mapping Concept 133
Hartanto Prawibowo, Rifky Ismail, Farika Tono Putri, Mohammad Tauviqirrahman, Robin Novriansyah, and Sulistia Nur

Classification of Stress Types Based on Electrodermal Activity Signal Using Multi-layer Perceptron Method 146
Clara Tyas Ratri, Osmalina Nur Rahma, Sayyidul Istighfar Ittaqillah, Rima Firda Alfisyah, Atin Asna Octavia, Fadli Ama, Khusnul Ain, and Rifai Chai

Efficient Gradient Boosting Integrated with Advanced Text Vectorization for Pulmonary Disease Classification 160
Phuoc-Hai Huynh

Sampling Rate Exploration to Obtain the Most Optimal Signal Resolution on 12-Lead ECG 172
Priyambada Cahya Nugraha, Anita Miftahul Maghfiroh, I Dewa Putu Arvian Danuartha, Muhammad Abyan Nafianto, Salma Syahidah Nurromadhona, Tetrik Fa'altin, and I Dewa Gede Hari Wisana

Increasing the Accuracy of Body Weight Sensor Readings Using the Kalman Filter on the Smart Baby Scale for Monitoring Toddler Nutrition . . . 184
Syaifudin, Nur Latifatul Arifah, and Endro Yulianto

Skin Cancer Identification Based on Dermoscopy Images Using Fractal and Gray Level Co-occurrence Matrix Features 202
Emmanuella Aurelia Rachel Passa, Endah Purwanti, Syifa Candiki Samatha, Alfian Pramudita Putra, and Rinci Kembang Hapsari

Shannon Entropy as Feature for Improved Prediction of Cardiac Risk Events in ECG Signals 219
Giovanny Andrés Piedrahita Solórzano, José Luis Ramírez Arias, and Daniel Botero Rosas

Facial Weakness Detection in Stroke Using Feedback Cascade Regression and Support Vector Machine 235
Mohamad Shavin, Tri Arief Sardjono, and Nada Fitrieyatul Hikmah

Multiclass Glaucoma Classification in Fundus Retina Images Using Modified U-Net Segmentation and VGG16 250
Andini Vira Salsabilla Zulkarnain Pakaya, Tri Arief Sardjono, Nada Fitrieyatul Hikmah, and Norma Hermawan

Utilizing IoT-Based Breath Analysis for Early Detection and Monitoring of Diabetes: A Deep Learning Approach 263
Nilakshi Maruti Mule and Dipti Durgesh Patil

Development of a Smart Punching Bag for Punch Force Level Identification Using Fuzzy Logic 279
Lilik Anifah, Nurhayati, Puput Wanarti Rusimamto, Muhamad Syarifuddien Zuhrie, and Haryanto

Optimation Environment Estimation for Maximum Fertility Level of Zea mays ssp. Plants in Automatic Fodder Using Fuzzy Logic 291
Haryanto, Dwi Kuswanto, Dian Neipa Purnamasari, and Lilik Anifah

A Monitoring System for Dissolved Oxygen and Temperature at Specific Water Depths Using Water Pressure Sensors 305
Muhammad Akmal Falih Rizqullah, Luthfi As-Sidqi, Mikhail Sannata Widjaya, Fiky Y. Suratman, Istiqomah, Eka Wardhani, and Kancitra Pharmawati

Development of a Bedside Monitor with an Embedded Digital Filter System to Reduce Noise in the Electrocardiogram Signal 319
Bambang Guruh Irianto, Anita Miftahul Maghfiroh, Reno Adi Surya Alsari, Abd. Kholiq, and Syevana Dita Musvika

Tool Wear Fault Detection and Classification Using Vibration Signal Analysis 336
Raja Siti Nur Adiimah Raja Aris, Fahmi Samsuri, Siti Nurshafina Zaharah Saffinee, Damhuji Rifai, Kharudin Ali, and Wasis Nugroho

PID Control System Design to Improve the Performance of Negative Pressure Wound Therapy 361
I. Dewa Gede Hari Wisana, Priyambada Cahya Nugraha, Levana Forra Wakidi, Danu Rogo, and Riqqah Dewiningrum

Design Simulation of a Voltage Balancing System in a Series Photovoltaic Array Based on an Isolated Bi-directional Converter with the Proportional-Integral Controller 378
Mukhammad Nur Hadi, Fahriza Balqis, Arya Kusumawardana, and Adi Izhar Che Ani

Is the Grid Ready? The Impact of Growing Electricity Demand on Low-Voltage Distribution 393
Abdul Halim Ikram Mohamed, Ruzlaini Ghoni, and Mohd Tarmizi Ibrahim

Design and Simulation of an Interleaved Boost Converter to Stabilize the Output Voltage under Varying Condition 402
Andre Dimas Yogatama, Febiola Sabrina, and Arya Kusumawardana

Drone Decentralized Scheduling for Rural Area Monitoring 418
Muhammad Zakiyullah Romdlony, Rashad Abul Khayr, Porman Pangaribuan, Yul Yunazwin Nazaruddin, and Tua Agustinus Tamba

Development of Pressure Control and Oxygen Control Methods for Continuous Mandatory Ventilation (CMV) and Synchronized Intermittent Mandatory Ventilation (SIMV) Modes in Ventilator 429
Basari and Sulthon Abdul Aziz

Evaluating Sand Resources for Radiation Protection: A Study of Silica and Iron Sand Distribution Along Yogyakarta Beach 446
Muhammad Fakhrurreza, Anisa Nur Istiqomah, and Rajendra Andita Nugraha

Development of a Non-Invasive Approach to Glucose, Cholesterol, and Hemoglobin Monitoring 459
Nur Hasanah Ahniar, Siddiq Andrianto, Rinda Nur Hidayati, Gita Rindang Lestari, and Anita Miftahul Maghfiroh

Detecting Major Depressive Disorder or Depression in Electroencephalogram Signal Using Long Short-Term Memory Networks 474
Mera Kartika Delimayanti, Euis Oktavianti, Bambang Warsuta, Rizky Adi, Umar Ali Ahmad, and Kenji Satou

Static Finite Element Analysis in 3D-Printed Insole with Thermoplastic Polyurethane (TPU) 75A for Individuals with Diabetics 485
Raden Achmad Candra Putra, Tri Riana Lestari, Muhammad I'Tikap, Syifa Fauziah, and Subyakto

Ten Hand Motion Classification Using Autoregression Feature Extraction Based EMG Signal 494
Triwiyanto, Endro Yulianto, Anita Miftahul Maghfiroh, and Wahyu Caesarendra

Feature Evaluation for Elbow Joint Angle Prediction Using EMG: A Preliminary Investigation 507
Triwiyanto, Levana Forra Wakidi, I. Putu Alit Pawana, Sari Luthfiyah, Bedjo Utomo, and Vijay Anant Athavale

ECG Digital Filter Processing: A Review 521
Triwiyanto, Triana Rahmawati, Endang Dian Setyoningsih, and Vugar Abdullayev

Author Index 541

Optimized Automated Standing Frame for Postural Therapy in Children with Cerebral Palsy

Johanes Wawan Joharwan[1(✉)], Ruzlaini Ghoni[2], Emanuel Budi Raharjo[1], Husein Arif Susanto[1], Ichsan Ramadhan[1], Damhuji Rifai[2], and Wasis Nugroho[2]

[1] Program Studi D3 Teknik Mesin, Jurusan Teknik Mesin, Sekolah Tinggi Teknologi "Warga" Surakarta, Jl. Raya-Baki Km. 2 Kwarasan, Sukoharjo, Jawa Tengah, Indonesia
joharwan@sttw.ac.id

[2] Faculty of Engineering Technology, University College TATI, Jalan Panchur, Teluk Kalong, Terengganu, Malaysia

Abstract. Cerebral palsy is a neurological condition that impairs muscle control and coordination, often preventing children from independently maintaining an upright standing position. Thus, to address this issue, static standing frames are commonly used as therapeutic aids, providing essential support and stability. However, existing frames fail to meet the dynamic needs of therapists and require manual adjustments. This study aims to develop an advanced, adjustable standing frame that meets the specific postural needs of children with cerebral palsy and their therapists. The frame features electric linear actuators that allow precise adjustments to standing position tilt, footrest angle, and the heights of lying and sitting positions, ensuring simplicity, safety, cleanliness, and accurate motion control. The backrest inclination can be customized using a reclining jog, and the leg tilt position is adjustable through secure connections, with lockable wheels enhancing mobility. The experiment involved five children aged 8–15 years, weighing between 12.6 and 20 kg and measuring between 105.2 and 117.5 cm in height. The frame consistently achieved a 28-s transition from lying to standing. This study contributes to the development of a highly functional and innovative static standing frame, addressing existing limitations. The frame's adjustability and user-friendliness significantly improve therapy sessions and enhance the overall therapeutic experience for both children with cerebral palsy and their therapists. The findings have significant implications for rehabilitation and posture management, promising improved outcomes for individuals with cerebral palsy.

Keywords: Static standing frame · Postural management · Cerebral palsy

1 Introduction

Cerebral palsy is a neurological condition that affects muscle control and coordination, resulting in difficulty in maintaining an independent upright standing position [1, 2]. To address this challenge, static standing frames are utilized as therapeutic aids, providing

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 1–11, 2026.
https://doi.org/10.1007/978-981-95-9233-3_1

necessary support and stability [3–19]. However, existing static standing frames have limitations that hinder meeting therapists' needs in facilitating movement abilities [8, 20–26]. Manual adjustments requiring therapist strength are necessary to alter the frame's position. Consequently, there is a pressing need to develop an improved static standing frame that addresses these limitations and meets the postural requirements of children with cerebral palsy and the needs of their therapists [27–39].

Despite ongoing research, there remains a gap in developing an adjustable static standing frame that fully addresses the needs of therapists and children with cerebral palsy, especially in automation [40–43]. Existing frames lack the necessary adjustability features and do not meet therapists' requirements to facilitate movement abilities in different positions [44–46]. Developing an advanced automated standing frame that caters to the specific postural requirements of children with cerebral palsy and meets therapists' needs is still unexplored [47–49]. This research aims to bridge this gap by designing an effective static standing apparatus tailored to the postural management needs of children with cerebral palsy and the requirements of their therapists.

This article presents a comprehensive study on the development of an advanced automated standing frame designed to optimize postural therapy in children with cerebral palsy. The research was conducted to address the limitations of existing standing frames, which often fail to meet the dynamic needs of both therapists and children, requiring manual adjustments and lacking flexibility. The primary objectives of this study are to create a standing frame that is adjustable, user-friendly, and capable of enhancing therapeutic outcomes. The methodology includes designing a static standing frame equipped with electric linear actuators to allow precise control over the standing position tilt, footrest angle, and heights of lying and sitting positions.

This study contributes in three significant ways: first, by filling the gap in existing research through the development of a highly functional and adjustable standing frame; second, by improving the therapy experience for both children with cerebral palsy and therapists; and third, by advancing rehabilitation practices and posture management techniques. The subsequent sections of the article detail the methodology, experimental results, and the broader implications of the study. The article concludes by emphasizing the significance of these contributions to enhancing rehabilitation outcomes and improving postural therapy sessions.

2 Materials and Method

The method used in this study is the product development processes, VDI 2221 [50] method in Fig. 1, which contains a systematic approach to developing system design and technical products. The purpose of this method is a general methodology for designing technical systems and products in support of methodical and systematic design, resulting in a more efficient work style. The method does not depend on the branch but discusses the content and design aspects of the organization and emphasizes broad applications in mechanical engineering, precision mechanics, change, software development, and process engineering planning. It is based on systems engineering and problem-solving.

The study involved 5 children diagnosed with cerebral palsy, aged between 8 and 15 years, selected to evaluate the therapeutic benefits of the optimized automated standing frame. All participants exhibited varying degrees of postural instability and limited independent movement; characteristics typical of children with cerebral palsy who benefit from postural support interventions. These children were selected through medical referrals based on their specific need for assisted-standing therapy, ensuring the study's relevance to the intended user group.

Data collection was conducted through structured observation and measurement. Each participant completed three sessions, with each session consisting of structured lying, sitting, and standing positions. The automated standing frame was adjusted according to the child's specific posture requirements, allowing for dynamic measurements of tilt, footrest angle, and backrest inclination. Throughout the sessions, postural alignment, stability, and comfort were monitored and recorded using both manual assessment and sensor data from the frame to ensure precise and consistent results.

An ANOVA test was employed to assess the standing frame's effectiveness in improving postural stability. This analysis allowed us to determine whether there were statistically significant improvements in postural alignment across the three positions (lying, sitting, and standing) over the course of the sessions. The ANOVA results indicated a significant positive impact on postural support, confirming the therapeutic benefits of the frame for the children in the study.

The Kinematic Equations for a static standing frame in a static scenario (no movement). Kinematic equations typically involve motion, relating displacement, velocity, and acceleration. The adjustability of the frame using mathematical notation for range and limits. The Standing Position Tilt is given as Eq. (1).

$$-\alpha \leq \theta \leq \beta \tag{1}$$

let θ (theta) represent the tilt angle of the standing position. We can define the range of motion as Eq. (2).

$$-\alpha \leq \theta \leq \beta \tag{2}$$

where α (alpha) is the minimum tilt angle (likely positive for a forward tilt) and β (beta) is the maximum tilt angle (likely negative for a backward tilt). The footrest angle equation is given as Eq. (3).

$$\gamma \leq \varphi \leq \delta \tag{3}$$

let φ (phi) represent the angle of the footrest, where γ (gamma) is the minimum footrest angle and δ (delta) is the maximum footrest angle. Meanwhile, the backrest inclination equation given in Eq. (4).

$$0^{\circ} \leq \psi \leq \theta_\max \tag{4}$$

let ψ (psi) represent the backrest inclination angle, where θ_max is the maximum backrest inclination angle (e.g., 45°).

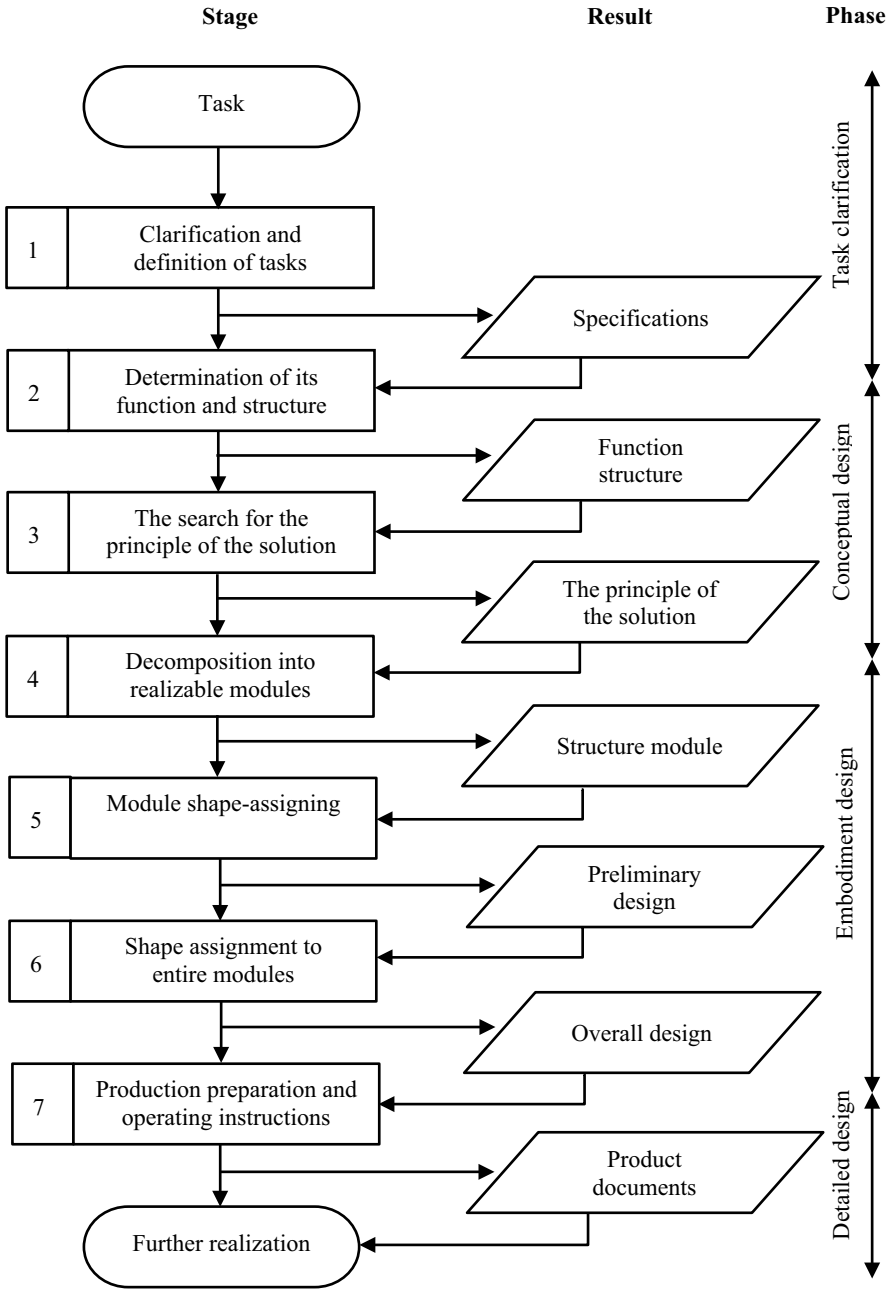


Fig. 1. A systematic method of VDI 2221

As part of the manufacture of products, the design process is divided into general stages of work, making the design approach transparent, rational, and independent of

a particular branch of the industry. The overall approach is divided into seven stages as illustrated in Fig. 1, which produce seven results. Depending on the task, all stages are completed or only a few, with stages repeated as necessary. In practice, individual stages are often combined into design phases, which help with the overall planning of the design process. Fig. 2 does not include the evaluation activities and decisions required at all stages, which determines whether it is necessary to repeat the previous stages. The Automated Standing Frame function is depicted in a block diagram of Fig. 2 that shows the relationship between input and output in the form of energy, material, and signal flows. The function of this static standing frame is for the postural management of children with cerebral palsy in lying down, sitting, and standing positions.

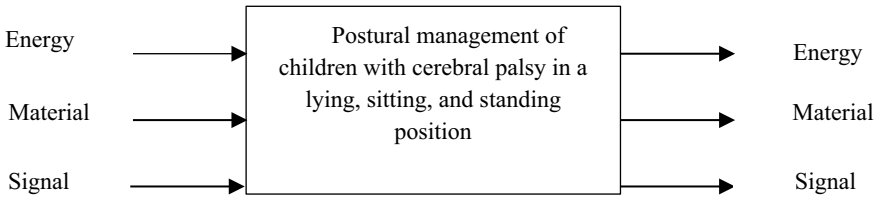


Fig. 2. A systematic method of VDI 2221

3 Result

The results show that a static standing frame developed according to the needs of therapists for the postural management of children with cerebral palsy in lying, sitting, and standing positions, as seen in Fig. 3. The static standing frame is adjustable with an electric linear actuator to adjust the tilt of the standing position, the tilt of the footrest, the height of the lying position, and the height of the sitting position according to the therapy needs. Electric linear actuators are used because of the simple, safe, clean, precise, and smooth motion control, in addition to the adjustable with a reclining jog backrest so that the slope of the back can be adjusted according to therapeutic needs. The tilt position of both legs can be adjusted with joints. The static standing frame also features a locking wheel for easy movement.

Stage 1 is necessary to clarify and define the task requirements requested by the customer or product planning department, including gathering all available information and finding the gaps, examining, and supplementing external requirements, adding specific company requirements, and defining and structuring tasks from the designer's point of view. The result is a specification (list of requirements) that can be created independently of any solution. Stage 2 determines the overall function and the most critical sub-functions (central functions). The classification and combination of these subfunctions into structures form the basis for finding solutions for the entire product or function. In Stage 3, a search was made for the principal solution for all subfunctions or initially for the most critical subfunctions of the function structure. Physical and other effects need to be selected for this purpose, and these should be able to be realized in principle by embodiment features. Phases 2 and 3 were included in the conceptual design

(a) (b) (c)

Fig. 3. Static standing frame position; a) Lying b) Sitting c) Standing

phase. In Stage 4, the principle solution is divided into realizable modules before beginning the complicated and time-consuming process of defining these modules in more concrete terms, which results in a module structure that, in contrast to the structure of functions or principle solutions, gives an early indication of the breakdown of solutions into groups and realizable elements (subsystems and system elements), which, together with their links (interfaces), are essential for their implementation. It can be represented in layout drawings, process flow diagrams, circuit diagrams, or software flow diagrams. The module structure is fundamental in the case of complex products, as it facilitates an efficient distribution of design efforts. It also helps with the identification and solution of embodiment design problems. So, for example, a distinction is made between design modules delimited according to the working principle: assembly modules, allowing easy assembly, maintenance modules, and recycling modules, and primary and variation modules, allowing modular product systems. Stage 5 consists of developing the module\key layout. The degree of refinement of geometry, materials, and other details should only be pursued to the extent that optimal design selection is possible. The result of this stage is a set of initial layouts for critical modules, which can be represented as scale drawings and circuit diagrams.

In Stage 6, the module's initial layout is supplemented by adding more detailed information about assemblies and components not previously included and combining all assemblies and components. It is often possible to define modules not included in Stage 5 by selecting standard or commercially available items. This stage produces a definitive layout that contains all the configuration information essential for product realization. The primary forms of representation are scale layout drawings, initial section lists, and instrumentation flowcharts. Stages 4 to 6 enter the embodiment design phase. In Stage 7, all final production, and operating instructions for which the design department is responsible were prepared. This stage thus overlaps with the previous one. The result of this stage is a set of product documents in the form of detailed assembly drawings, parts lists, and production, assembly, testing, transportation, and operation instructions. Phase 7 enters the detailed design phase.

An experiment was performed on children with cerebral palsy to see the duration from lying to standing position with children of age range from 8 to 15 years old and weight of 11.5 to 20 kg. The experiment abides by the Nuremberg Code of Ethics of medical research. The result shows the consistent time to accomplish, thus easing the therapist's task. The result is shown in Table 1.

Table 1. Duration from lying to standing position of cerebral palsy children.

Gender	Age	Height (cm)	Weight (Kg)	Time (s)
Male	10	117.5	20	28
Male	9	111.3	12.6	28
Male	8	105.2	17.2	28
Female	15	115.5	11.5	28
Female	9	111	13	28

4 Discussion

The results demonstrated the successful development of an advanced automated standing frame for optimal postural therapy in children with cerebral palsy, thus addressing the limitations of existing static standing frames, which require manual adjustments and do not fully meet therapists' needs. The designed frame, incorporating automation and adjustability features, provides support, stability, and precise control over various parameters, including tilt, footrest angle, and sitting and lying positions. This innovative approach ensures improved therapy sessions and enhances the overall experience for therapists and children with cerebral palsy.

Compared to other studies, this research focuses on developing an advanced automated standing frame to meet the postural requirements of children with cerebral palsy and address therapists' needs. Previous studies have acknowledged the limitations of existing frames but have not explored the integration of automation and adjustability features. Our study presents a unique and effective solution to optimize postural therapy by incorporating electric linear actuators and a tailored design. This innovative approach enhances safety, precision, and ease of use, providing a valuable tool for therapists working with children with cerebral palsy.

While this study brings significant advancements, it also has some limitations. Firstly, the sample size was relatively small, consisting of only five children with cerebral palsy. A larger sample size would increase the study's statistical power and improve generalizability. Additionally, the study focused on a specific age range and weight distribution, which may limit the applicability of the frame to a broader population. Future research should involve a more diverse participant pool to validate the effectiveness and applicability of the advanced automated standing frame across a wider range of ages and weight categories.

The findings of this study have significant implications for therapy practice and the broader field of cerebral palsy management. The advanced automated standing frame developed in this research addresses the limitations of manual adjustment frames, offering enhanced adjustability and precise control that contribute to improved rehabilitation outcomes and postural management for children with cerebral palsy. These results are consistent with other studies that emphasize the importance of tailored therapeutic devices in enhancing patient outcomes. However, one limitation of this study is the relatively small sample size, which may affect the generalizability of the findings. Additionally, while the frame's automation and adjustability offer clear benefits, the reliance on technology may present challenges in settings with limited resources. Despite these limitations, the study opens new avenues for further advancements in automated standing frames and encourages the integration of technology to optimize therapy approaches. The implications of this research extend beyond individual therapy sessions, potentially benefiting society by improving the quality of life for children with cerebral palsy and easing the workload for therapists. Ultimately, these findings have the potential to positively impact the lives of both therapists and children, leading to better overall therapy experiences and outcomes.

5 Conclusion

This study aimed to design and develop an advanced automated standing frame for optimal postural therapy in children with cerebral palsy, addressing the limitations of existing static frames. The findings demonstrated the successful development of an adjustable static standing frame equipped with electric linear actuators, which provided consistent and reliable support for postural therapy. The frame achieved a transition time of 28 s from lying to standing, indicating efficient performance. Quantitatively, the frame's adjustability was shown to significantly enhance therapy sessions, as evidenced by improved ease of use and control during experiments. Future work should include larger-scale studies with a diverse population to further analyze the frame's effectiveness. Additionally, exploring long-term effects on postural development and functional abilities, as well as integrating sensors and data analytics for enhanced functionality and personalized therapy, will be crucial. Continuous research in this area holds great potential for advancing postural therapy and significantly improving the quality of life for individuals with cerebral palsy.

References

1. Vitrikas K, Dalton H, Breish D (2020) Cerebral palsy: an overview. https://doi.org/10.4274/hamidiyemedj.galenos.2021.72792
2. Upadhyay J, Tiwari N, Ansari MN (2020) Cerebral palsy: aetiology, pathophysiology and therapeutic interventions. Clin Exp Pharmacol Physiol 47. https://doi.org/10.1111/1440-1681.13379
3. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB, Livingstone C (2001) Evidence-based medicine: how to practice and teach EBM. J Intensive Care Med 16:155–156
4. Vignos PJ, Wagner MB, Karlinchak B, Katirji B (1996) Evaluation of a program for long-term treatment of Duchenne muscular dystrophy. J Bone Joint Surg 78-A:1844–1852

5. Baker K, Cassidy E, Rone-Adams S (2007) Therapeutic standing for people with multiple sclerosis efficacy and feasibility. Int J Ther Rehabil 14:104–110
6. Tremblay F, Malouin F, Carol L, Richards DF (1990) Effects of prolonged muscle stretch on reflex and voluntary muscle activations in children with spastic cerebral palsy. Scand J Rehab Med 22:171–180
7. Walter JS, Sola PG, Sacks J, Lucero Y, Langbein E, Weaver F (1999) Indications for a home standing program for individuals with spinal cord injury. J Spinal Cord Med 22:152–158. https://doi.org/10.1080/10790268.1999.11719564
8. Stuberg WA (1992) Considerations related to weight-bearing programs in children with developmental disabilities. Phys Ther 72:1–6
9. Macias-Merlo L, Bagur-Calafat C, Girabent-Farrés MA, Stuberg W (2016) Effects of the standing program with hip abduction on hip acetabular development in children with spastic diplegia cerebral palsy. Disabil Rehabil 38:1075–1081. https://doi.org/10.3109/09638288.2015.1100221
10. Stein TP, Gaprindashvili T (1994) Spaceflight and protein metabolism, with special reference to humans14. Spacelight Protein Metab 60:8065–8195
11. Henderson RC, Kairalla J, Stevenson RD (2004) Predicting low bone density in children and young adults with quadriplegic cerebral palsy. Dev Med Child Neurol 46:416–419
12. Ward K, Alsop C, Caulton J, Rubin C, Adams J, Mughal Z (2004) Low magnitude mechanical loading is osteogenic in children with disabling conditions. J Bone Miner Res 19:360–369. https://doi.org/10.1359/JBMR.040129
13. Dunn RB et al (1998) Follow-up assessment of standing mobility device users. Assist Technol 10:84–93. https://doi.org/10.1080/10400435.1998.10131966
14. Eng JJ, Levins SM, Townson AF, Mah-Jones D, Bremner J (2001) Grant Huston: use of prolonged standing for individuals with spinal cord injuries. Phys Ther 81:1392–1399
15. Ahlborg L, Andersson C, Julin P (2006) Whole-body vibration training compared with resistance training: effect on spasticity, muscle strength and motor performance in adults with cerebral palsy. J Rehabil Med 38:302–308. https://doi.org/10.1080/16501970600680262
16. Netz Y et al (2007) Use of a device to support standing during a physical activity program to improve function of individuals with disabilities who reside in a nursing home. Disabil Rehabil Assist Technol 2:43–49. https://doi.org/10.1080/17483100601143371
17. Cheng HYK, Ju YY, Chen CL, Chuang LL, Cheng CH (2015) Effects of whole body vibration on spasticity and lower extremity function in children with cerebral palsy. Hum Mov Sci 39:65–72. https://doi.org/10.1016/j.humov.2014.11.003
18. Edwards LC, Layne CS (2007) Effect of dynamic weight bearing on neuromuscular activation after spinal cord injury. Am J Phys Med Rehabil 86:499–506. https://doi.org/10.1097/PHM.0b013e31805b764b
19. Arva J et al (2009) RESNA position on the application of wheelchair standing devices. Assist Technol 21:161–168. https://doi.org/10.1080/10400430903175622
20. Carneiro T, Oliveira J, Baptista AJ, de Castro PMST (2022) Enhancement of the lean DfX product development framework and application to the design of an AGV structure. Designs (Basel) 6. https://doi.org/10.3390/designs6060114
21. Tomiyama T, Gu P, Jin Y, Lutters D, Kind C, Kimura F (2009) Design methodologies: industrial and educational applications. CIRP Ann Manuf Technol 58:543–565. https://doi.org/10.1016/j.cirp.2009.09.003
22. Zschiebsch W, Filippatos A, Bohm R (2022) A digital-based design methodology for the optimization of high-performance multi-material structures. IOP Conf Ser Mater Sci Eng 1226:012078. https://doi.org/10.1088/1757-899x/1226/1/012078
23. Birkhofer H, Jänsch J, Kloberdanz H (2005) An extensive and detailed view of the application of design methods and methodology in industry. In: International conference on engineering design, pp 1–12

24. Forster J, Cartmell M, Fothergill P (1990) Artificial Intelligence Techniques. Des Methodol Mech Eng:819–827
25. Bucherta T, Kaluza A, Halstenberg FA, Lindow K, Hayka H, Stark R (2014) Enabling product development engineers to select and combine methods for sustainable design. In: Procedia CIRP. Elsevier B.V, pp 413–418. https://doi.org/10.1016/j.procir.2014.06.025
26. Sackett, D.L., Straus, S.E., Richardson, W.S., Rosenberg, W., Haynes, R.B., Livingstone, C.: Evidence-based medicine: how to practice and teach EBM, 2nd Edition. (2001).
27. Graham HK et al (2016) Cerebral palsy. Nat Rev Dis Primers 2. https://doi.org/10.1038/nrdp.2015.82
28. Goodwin J et al (2018) Understanding frames: a qualitative study of young people's experiences of using standing frames as part of postural management for cerebral palsy. Child Care Health Dev 44:203–211. https://doi.org/10.1111/cch.12540
29. Saavedra SL, Goodworth AD (2019) Postural control in children and youth with cerebral palsy. In: Cerebral palsy. Springer International Publishing, pp 1–21. https://doi.org/10.1007/978-3-319-50592-3_161-1
30. Pawlyn J, Carnaby S (2009) Profound intellectual and multiple disabilities: nursing complex needs. Wiley-Blackwell
31. Pierret J, Caudron S, Paysant J, Beyaert C (2021) Impaired postural control of axial segments in children with cerebral palsy. Gait Posture 86:266–272. https://doi.org/10.1016/j.gaitpost.2021.03.012
32. Livingstone RW, Paleg GS (2023) Use of Overground supported-stepping devices for non-ambulant children, adolescents, and adults with cerebral palsy: a scoping review. Disabilities 3:165–195. https://doi.org/10.3390/disabilities3020012
33. Ekström Ahl L et al (2005) Functional therapy for children with cerebral palsy: an ecological approach. Dev Med Child Neurol 47:613–619
34. Machado J, Seabra E (2011) Design of an adapted standing frame for rehabilitation of children with mental deficiency. The Romanian review precision mechanics
35. Rosenbaum PL et al (2002) Prognosis for gross motor function in cerebral palsy creation of motor development curves. JAMA 22:1357–1363
36. Saether R, Helbostad JL, Riphagen II, Vik T (2013) Clinical tools to assess balance in children and adults with cerebral palsy: a systematic review. Dev Med Child Neurol 55:988–999. https://doi.org/10.1111/dmcn.12162
37. Ibrahim F, Osman NA (2007) A new modification of adjustable standing Frame for children with cerebral palsy. In: IFMBE proceedings. pp 167–170
38. Sibley KM, Beauchamp MK, Van Ooteghem K, Paterson M, Wittmeier KD (2017) Components of standing postural control evaluated in Pediatric balance measures: a scoping review. Arch Phys Med Rehabil 98:2066–2078.e4. https://doi.org/10.1016/j.apmr.2017.02.032
39. Stetter, R., Simundsson, A.: Systematic product development of control and diagnosis functionalities. J Phys Conf Ser . Institute of Physics Publishing (2017). doi:https://doi.org/10.1088/1742-6596/783/1/012063.
40. Eigner M, Dickopf T, Apostolov H (2017) The evolution of the V-Model: From VDI 2206 to a system engineering based approach for developing cybertronic systems. pp 382–393. https://doi.org/10.1007/978-3-319-72905-3_34.
41. Brogren E, Hadders-Algra M, Forssberg H (1998) Postural control in sitting children with cerebral palsy. Neurosci Biobehav Rev 22:591–596
42. Vielhaber M (2011) Design to knowledge—a root design principle. In: 56th International scientific colloquium
43. Eder WE (1993) A European outlook on engineering design in education. J Eng Educ 82:118–122. https://doi.org/10.1002/j.2168-9830.1993.tb00085.x

44. Goodwin J et al (2018) Understanding frames: a UK survey of parents and professionals regarding the use of standing frames for children with cerebral palsy. Child Care Health Dev 44:195–202. https://doi.org/10.1111/CCH.12505
45. Gibson SK, Sprod JA, Maher CA (2009) The use of standing frames for contracture management for nonmobile children with cerebral palsy. Int J Rehabil Res 32:316–323. https://doi.org/10.1097/MRR.0B013E32831E4501
46. Goodwin J et al (2018) Standing frames for children with cerebral palsy: a mixed-methods feasibility study. Health Technol Assess (Rockv) 22:1–231. https://doi.org/10.3310/HTA22500
47. Daniels N, Pinnington L, Clift M, Gopsill C, Armstrong J, Ward C (2005) A review of standing frame features and their impact on user acceptability. Int J Ther Rehabil 12. https://doi.org/10.12968/ijtr.2005.12.8.19539
48. Paleg GS, Smith BA, Glickman LB (2013) Systematic review and evidence-based clinical recommendations for dosing of pediatric supported standing programs. https://doi.org/10.1097/PEP.0b013e318299d5e7
49. Laranjeira M et al (2018) Mechanical Design of a Standing Frame adapted for children with mental deficiency. Procedia CIRP. https://doi.org/10.1016/j.procir.2018.04.086
50. Jänsch J, Birkhofer H (2006) The development of the guideline VDI 2221—the change of direction. In: 9th International design conference. Design 2006

How the Entrepreneurial Ecosystem Drives Success in Telemedicine Healthcare Businesses: A Systematic Review

Puan Ayu Maharani, Mardhatillah Shanti(✉), and Agung Purnomo

BINUS Business School Undergraduate Program, Entrepreneurship Department, Bina Nusantara University, Jakarta 11480, Indonesia
mardhatillah.shanti@binus.edu

Abstract. The growing demand for telemedicine services, accelerated by the COVID-19 pandemic and new normal, has highlighted the critical roles of entrepreneurial ecosystems in fostering innovation, improving healthcare access, and promising sustainability of digital health ventures. This study systematically reviews the entrepreneurial ecosystem in telemedicine healthcare businesses, focusing on the elements and roles contributing to their success. This study introduces a novel comprehensive framework of the entrepreneurial ecosystem in telemedicine, addressing gaps in previous research that focused narrowly on technological, demographic, or operational aspects, thereby providing a holistic understanding of how these factors collectively drive innovation and sustainability in telemedicine healthcare businesses. A comprehensive review of literature from the Scopus database was conducted by applying the PRISMA protocols and the Gioia method. Based on the research findings, this review proposes the 4Ps fishbone conceptual framework of the entrepreneurial ecosystem that drives success in telemedicine healthcare businesses. This review identified key elements of the entrepreneurial ecosystem, such as formal institutions, entrepreneurship culture, and networks that significantly influence the growth and sustainability of telemedicine ventures. The prominent roles of the entrepreneurial ecosystem in the telemedicine health service business include people development, process optimization, policy framework, and platform advancement. This study highlights how the entrepreneurial ecosystem drives success in telemedicine healthcare businesses, providing strategic insights for entrepreneurship, practical guidance for managers, and policy recommendations to foster growth.

Keywords: Digital Business · Entrepreneurship · Entrepreneurial Ecosystem · Healthcare Business · Telemedicine

1 Introduction

The study of telemedicine has rapidly evolved, propelled by technological innovations and the growing necessity for affordable health services. Research has shown a steady increase in telemedicine applications, particularly in the background of mobile health

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 12–26, 2026.
https://doi.org/10.1007/978-981-95-9233-3_2

(mHealth) and digital health innovations. The COVID-19 pandemic and new normal further accelerated this trend, highlighting the potential of telemedicine to alleviate pressure on traditional healthcare systems by providing remote consultations and monitoring [1, 2]. Studies have emphasized the importance of integrating telehealth into healthcare practices to ensure continuity of care, especially in rural and underserved areas [3, 4].

Entrepreneurship in the health sector has also seen significant growth, with social enterprises playing a crucial role in increasing access to primary healthcare in low-resource settings [5]. These ventures often adopt innovative business models to address social problems while maintaining financial sustainability [6, 7]. The role of digital health technologies in entrepreneurship has been highlighted, particularly in improving health outcomes and ensuring the scalability of health initiatives [8, 9]. The intersection of health and entrepreneurship is pivotal in addressing global health challenges and fostering innovation.

The growing reliance on digital health solutions underscores the urgency of studying the entrepreneurship ecosystem in telemedicine. As telemedicine becomes more integral to healthcare delivery, understanding the entrepreneurial ecosystem that supports these innovations is critical. This includes examining the roles of various stakeholders, the challenges they face, and the strategies that lead to successful telehealth ventures [1, 4]. A robust entrepreneurship ecosystem is essential for telemedicine's sustained growth and impact [3, 10].

Applying a systematic literature review (SLR) methodology to research the entrepreneurship ecosystem in telemedicine offers several advantages. SLR allows for a comprehensive and structured literature analysis, identifying patterns, gaps, and trends [6, 11]. This method ensures a rigorous evaluation of the evidence base, providing valuable insights into the factors that influence the success of telehealth ventures [12, 13]. By synthesizing diverse studies, SLR can highlight best practices and areas requiring further investigation.

The entrepreneurial ecosystem significantly influences the success of telemedicine healthcare businesses, as evidenced by various studies employing diverse methodologies. Alzagladi's systematic review identifies key facilitators, such as patient engagement and regulatory changes, but it does not comprehensively analyze the entrepreneurial ecosystem's broader context [14]. Mohammadzadeh et al. focus on telemedicine's roles in underserved populations, highlighting operational challenges, but they do not address systemic entrepreneurial factors [15]. Waseem et al. focused on disparities in telemedicine success among thoracic cancer patients, revealing that sociodemographic factors significantly impact outcomes. Yet, this study lacked a comprehensive view of the entrepreneurial ecosystem [16]. Mabeza et al. provide insights into chronic disease management through synchronous telemedicine, but their findings are limited to specific conditions, missing a holistic view of the entrepreneurship ecosystem [17]. Kia et al. conducted a meta-analysis on emergency department telemedicine, which, while robust, does not explore entrepreneurial dynamics [18]. Ftouni et al. highlight training needs during the pandemic but do not connect these to the entrepreneurial ecosystem [19].

Despite these valuable contributions, a comprehensive understanding of the entrepreneurship ecosystem in telemedicine remains underexplored. There is a research gap in synthesizing the diverse findings into a cohesive framework that addresses the roles within the ecosystem. Comprehensive research on the entrepreneurial ecosystem in telemedicine is limited, hindering the comprehension of critical factors for successful ventures. This research gap can be filled by employing a systematic literature review (SLR) methodology and grounded theory of the Gioia method to synthesize these elements, offering a comprehensive understanding of how the entrepreneurial ecosystem can drive success in telemedicine healthcare businesses [12, 13, 20]. The research question in this review is how the literature development status is related to the elements and role in studying the entrepreneurship ecosystem in telemedicine worldwide. Therefore, this study aims to determine the literature development status related to the elements and role in studying the entrepreneurship ecosystem in telemedicine healthcare businesses worldwide. This study is essential as it fills the research gap by explaining how the entrepreneurial ecosystem supports telemedicine success, benefiting society through improved healthcare access, innovation, and scalable digital health solutions.

2 Materials and Method

This work has carried out a Systematic Literature Review (SLR) following the PRISMA protocol. The PRISMA source is chosen to be followed as it is performed within a rigorous framework, ensuring a transparent, systematic review [21, 22]. It also requires guidelines for selecting, eligibility, and exclusion articles, which should provide a full-scale covering of the relevant literature concerning the purpose of the research that is crucial for synthesizing evidence effectively [23]. The study collected data from the Scopus Database, which holds telemedicine and entrepreneurship ecosystems. The coverage of Scopus is of high quality and is broad, further holding relevance for this review [24].

The research questions (RQs) guiding the study were structured as follows.

RQ1: What are the elements of the entrepreneurial ecosystem that drive success in the telemedicine healthcare business?

RQ2: What are the roles of the entrepreneurial ecosystem that drive success in the telemedicine healthcare business?

RQ3: What are the key findings of the entrepreneurial ecosystem that drive success in the telemedicine healthcare business?

The search command keywords in the document search have been “telemedicine” and “entrepreneurship,” combined with the corresponding terms. The search query option used in data mining was as follows TITLE-ABS-KEY (telemedicine* AND entrepreneur*) AND PUBYEAR >2012 AND PUBYEAR <2024 AND (LIMIT-TO (DOCTYPE, “ar”)) as of July 2024. Twenty-four documents were reviewed and met the inclusion criteria for telemedicine and entrepreneurship, publications in journals, and articles up to 2023, as seen in Table 1.

Table 1. Inclusion criteria.

Inclusion criteria	Description
Topics	Address telemedicine and entrepreneurship
Document type	Article
Publication year	Published up to the year 2023

For the analysis, RQ1 and RQ2 have been analyzed descriptively and thematically using the Gioia method, which offers a structured approach to analyzing qualitative data, enabling a deeper understanding of the elements and role identified of entrepreneurship to enhance the development of telemedicine. The Gioia method provides a systematic framework for organizing and interpreting textual data, facilitating the extraction of meaningful insights from the literature [20]. RQ3 has been analyzed descriptively to synergize the main findings of elements and roles in the study of entrepreneurial ecosystems in telemedicine visualized in the conceptual framework. The PRISMA Protocols can be seen in Fig. 1.

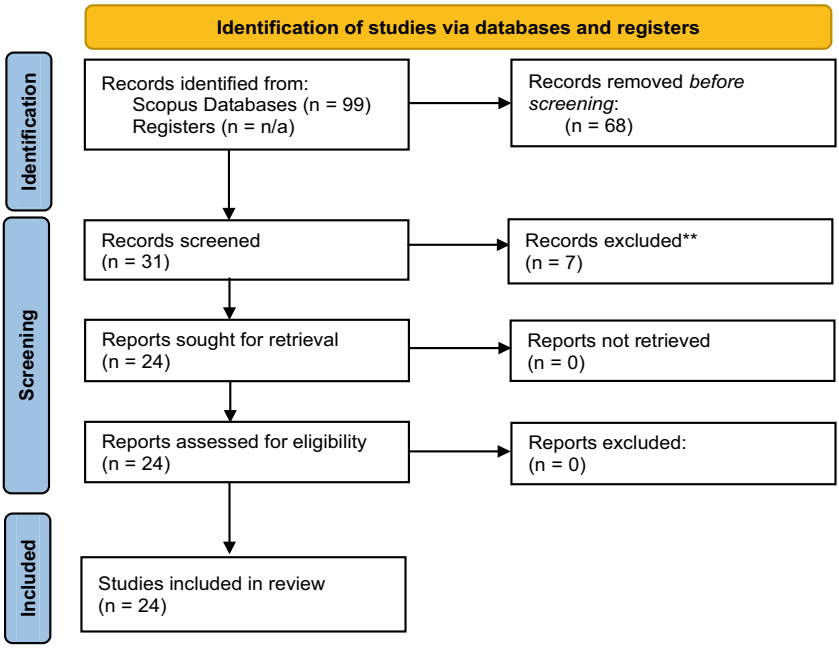


Fig. 1. PRISMA protocols

3 Result

3.1 Elements of Entrepreneurial Ecosystem in Telemedicine Healthcare Business

Table 2 presents a refined version of the elements in the entrepreneurial ecosystem within telemedicine research. Less relevant or overly specific sections have been condensed to streamline the findings, ensuring clarity and focus. Despite this reduction, the number of authors listed remains unchanged to maintain the integrity of the original contributions.

The entrepreneurial ecosystem in telemedicine relies on formal solid institutions, which include the involvement of healthcare professionals [1], supportive policies [25], regulation management capacity [8], the Health Professions Council of South Africa [3], and legislative efforts framing telemedicine as an integral [26]. Entrepreneurship culture within the telemedicine ecosystem is driven by a culture of innovation [25], fostering an entrepreneurial mindset [25], social trust [27], entrepreneurial creativity [28], and digital health entrepreneurship [2]. The physical infrastructure essential to the telemedicine entrepreneurial ecosystem includes mobile health (mHealth) technologies [1], technology integration [25], robust technological infrastructure [8], digital technologies [4], and telemedicine infrastructure [29]. The entrepreneurial ecosystem's demand for telemedicine is fueled by raising public recognition [1], increasing market demand [8], telehealth service [3], e-commerce [11], and telemedicine sustains healthcare access during crises [26]. Networks within the telemedicine entrepreneurial ecosystem thrive on connected health solutions [1], global networking efforts [25], strategic partnerships [30], health-tech ecosystem [8], and collaboration by UN agencies [4].

The leadership within the entrepreneurial ecosystem for telemedicine includes leadership development initiatives [25], public-private engagement efforts [31], the role of NHS Trusts [27], policy entrepreneurs within the Ministry of Health [7], and healthcare managers focusing on efficiency [9], which are crucial for driving innovation. Talent within the telemedicine entrepreneurial ecosystem focuses on disease management [1], essential resources [32], continuous professional development for healthcare professionals [4], m-health education [2], and technology learning curve for employees [33]. The finance component of the telemedicine entrepreneurial ecosystem is driven by startup entrepreneurship [1], access to essential resources [25], the ability to generate revenue through innovation [8], health insurance investment banking [3], and access to external financial capital [33]. New knowledge within the telemedicine entrepreneurial ecosystem is continuously developed through gamification elements [1], mentorship opportunities [25], health innovations [4], education and training services [27], and key activities [32]. The entrepreneurial ecosystem's intermediate services in telemedicine include a supportive environment [25], advisory services [27], strong partnerships with healthcare platforms [3], high-quality service delivery [28], and coordination of the supply of essential products [10]. Lastly, productive entrepreneurship in telemedicine focuses on telemedicine franchising [6], enhancing service value [8], the adoption of remote care systems [34], healthcare delivery [8], and health monitoring [8].

Table 2. Elements of entrepreneurial ecosystem in telemedicine

No.	Elements of entrepreneurial ecosystem	Elements of entrepreneurial ecosystem in telemedicine
1	Formal institutions	Healthcare professionals' involvement [1], policy support [25], regulation management capacity [8], Health Professions Council of South Africa [3], and legislative efforts framing telemedicine as integral [26]
2	Entrepreneurship culture	**Innovation Culture** [25]**, Entrepreneurial Mindset** [25]**, Social trust** [27]**, Entrepreneurial creativity** [28]**, and Digital health entrepreneurship** [2]
3	Physical infrastructure	Mobile health (health) [1], Technology Integration [25], Robust technological infrastructure [8], Digital technologies [4], and Telemedicine infrastructure [29]
4	Demand	Raising public recognition [1], Market demand [8], Telehealth services [3], E-commerce [11], and Telemedicine sustaining healthcare access during crises [26]
5	Networks	Connected health [1], Global Networking [25], Strategic partnerships [30], Health-tech ecosystem [8], Collaboration by UN agencies [4]
6	Leadership	Leadership Development [25], Public-private engagement [31], Role of NHS Trusts [27], Policy entrepreneurs within the Ministry of Health [7], Healthcare managers focusing on efficiency [9]
7	Talent	Disease management [1], Key resources [32], Training and guidance for health professionals [4], M-health education [2], and Technology learning curve for employees [33]
8	Finance	Startups and individual entrepreneurs [1], Access to Resources [25], Revenue generation ability [8], Health insurance investment banking [3], and access to external financial capital [33]

(*continued*)

Table 2. (*continued*)

No.	Elements of entrepreneurial ecosystem	Elements of entrepreneurial ecosystem in telemedicine
9	New knowledge	Gamification elements [1], Mentorship Opportunities [25], Health innovations [4], Education and training services [27], Key activities [32]
10	Intermediate services	Supportive Environment [25], Advisory services [27], Partnership with Healthtap Inc. [3], High-quality service delivery [28] and coordination of the supply of essential products [10]
11	Productive entrepreneurship	Telemedicine franchising and mobile units [6], Service value and effectiveness [8], Adoption of remote care systems [34], Healthcare delivery [8], and Health monitoring [8]

This study identifies critical elements of the entrepreneurial ecosystem that drive success in telemedicine healthcare businesses, including formal institutions, entrepreneurial culture, networks, leadership, talent, and finance. This comprehensive framework fits with contemporary literature that underscores the multifaceted nature of telemedicine adoption and success, particularly the role of formal institutions and networks in fostering a supportive environment for telemedicine [35]. It also underscores the need for a robust business framework incorporating stakeholder perspectives, resonating with the emphasis on leadership and talent as crucial ecosystem components [32]. Unlike studies focusing on technology, my research offers a holistic view of ecosystem dynamics, addressing gaps and providing a more nuanced understanding of how these factors collectively enhance telemedicine scalability and innovation [36].

3.2 Roles of the Entrepreneurial Ecosystem in Telemedicine Healthcare Business

Following the data structuring and analysis process using the Gioia method, four aggregate dimensions were developed to describe the roles faced when implementing the entrepreneurship ecosystem in telemedicine, as shown in Fig. 2. These roles' aggregate dimensions are people development, process optimization, policy framework, and platform advancement. Each aggregate dimension is explained in the following section.

People Development.
The findings of this research reveal that the application of an entrepreneurial ecosystem in telemedicine has a role of professional involvement. The application of the entrepreneurial ecosystem in telemedicine involves health workers [1], the entrepreneurial ecosystem plays a vital role in driving innovation [25], and international organizations play an essential role in supporting social entrepreneurship [6].

Another role finding is stress management and support. Implementing the entrepreneurship ecosystem in telemedicine via telephone as a supportive intervention to improve stress prevention programs [37]. Another role discovery is building trust. The application of the entrepreneurial ecosystem in telemedicine has a vital role in social trust for the development and effectiveness of the telemedicine health services market [27].

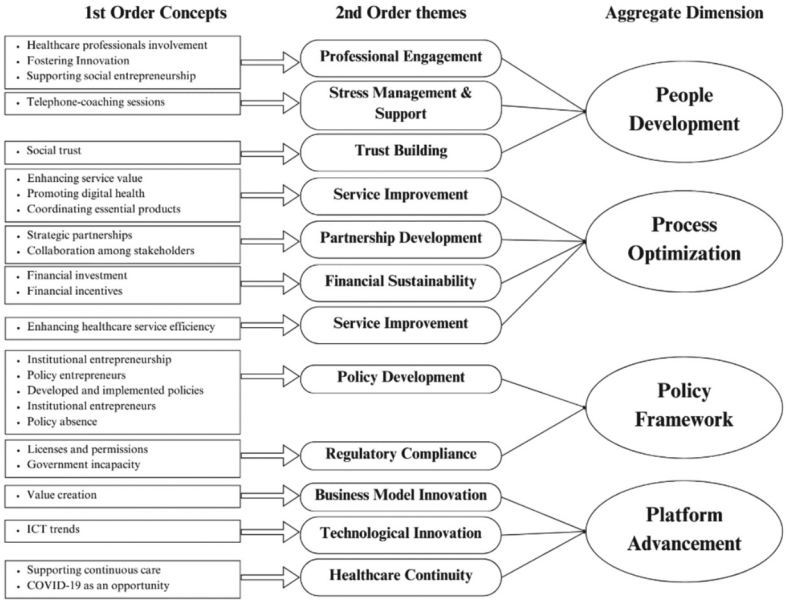


Fig. 2. Roles of the entrepreneurial ecosystem in telemedicine healthcare business

Process Optimisation.
The findings of this research reveal that the application of the entrepreneurship ecosystem in telemedicine has a role in improving services. Implementation of an entrepreneurial ecosystem in telemedicine plays a role in increasing service value [8], promoting digital health through the application of the disconnect platform [3, 38], and coordinating the supply of essential products to hospitals [10]. Another role finding is partnership development. Implementing the entrepreneurship ecosystem in telemedicine involves strategic partnerships [33] and collaboration between various stakeholders [9]. Another role finding is financial sustainability. Implementing an entrepreneurial ecosystem in telemedicine has a significant financial investment role for telemedicine systems [33], and economic incentives are substantial for hospitals and health systems [34]. Another role discovery is service improvement. Implementing the entrepreneurial ecosystem in telemedicine has a role in increasing the efficiency of health services [26].

Policy Framework.
The findings of this research reveal that the implementation of the entrepreneurial ecosystem in telemedicine has a role in policy development. Implementation of the

entrepreneurial ecosystem in telemedicine has the role of institutional entrepreneurship [31], policy entrepreneurs [7], developed and implemented policies [13], institutional entrepreneurs [31], and policy absence [39]. Another role finding is regulatory compliance. Implementing the entrepreneurial ecosystem in telemedicine involves licenses and permissions [32] and government incapacity [40].

Platform Advancement.
The findings of this research reveal that the implementation of the entrepreneurial ecosystem in telemedicine has the role of business model innovation. Implementing the entrepreneurial ecosystem in telemedicine has a value-creation role [11]. Another role found is technological innovation. Implementing the entrepreneurial ecosystem in telemedicine has the role of ICT trends [28]. Another role finding is healthcare continuity. Implementing the entrepreneurial ecosystem in telemedicine supports continuous care [4] and COVID-19 as an opportunity [2].

3.3 Key Findings

This research has identified that the entrepreneurial ecosystem drives success in telemedicine healthcare businesses through eleven main elements: formal institutions, entrepreneurial culture, demand, physical infrastructure, networks, talent, leadership, finance, new knowledge, intermediary services, and productive entrepreneurship. The systematic review reveals that the entrepreneurial ecosystem drives success in telemedicine healthcare businesses through four prominent roles: people development, process optimization, policy framework, and platform advancement. Therefore, this study proposes and visualizes the 4Ps fishbone conceptual framework as the principal elements and roles of the entrepreneurial ecosystem that drive success in telemedicine healthcare businesses, as seen in Fig. 3.

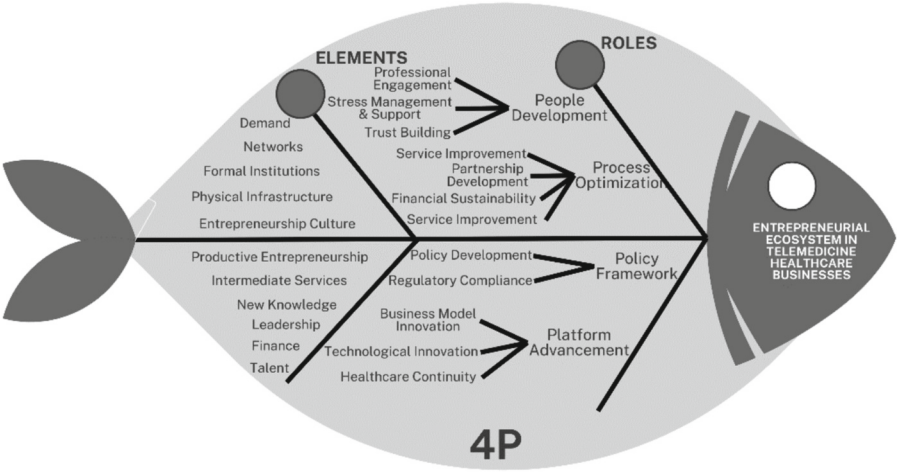


Fig. 3. The 4P fishbone conceptual framework of main elements and roles

4 Discussion

The systematic review of the entrepreneurial ecosystem in telemedicine healthcare businesses reveals that critical elements such as formal institutions, demand, entrepreneurial culture, networks, leadership, talent, finance, physical infrastructure, intermediary services, new knowledge, and productive entrepreneurship significantly contribute to their success. This aligns with findings from previous studies, which underscore the significance of entrepreneurial competencies and culture in fostering business growth. For instance, Gunartin et al. highlight that entrepreneurial competencies are foundational for success in SMEs, reinforcing the notion that a robust entrepreneurial culture is essential for innovation and scalability in telemedicine [41]. Similarly, Nti's research underscores the role of strategic orientation and resource management in entrepreneurial success, which resonates with the identified elements in our study [42]. Furthermore, the role of networks and institutional support, as discussed by Wahyuningtyas et al., is crucial in enhancing entrepreneurial capabilities within the ecosystem [43]. This systematic review addresses gaps in the existing literature by providing a comprehensive analysis of how these elements interact specifically within the telemedicine sector. This area has not been extensively explored in prior studies.

The analysis of the roles of the entrepreneurial ecosystem in telemedicine healthcare businesses reveals that crucial elements such as people development, process optimization, policy framework, and platform advancement are critical for success. This aligns with findings from prior research that underscore the significance of healthcare professionals' attitudes and training in integrating telemedicine effectively into healthcare systems [44]. For instance, Wernhart et al. highlight the dispositions of healthcare personnel are essential for successful eHealth connection [45], while Těšinová points out that community involvement significantly influences telemedicine adoption [46]. Furthermore, the systematic review underscores that a robust policy framework is necessary to address barriers and enhance the operational capacity of telemedicine services, echoing the sentiments of Alami et al., who noted the impact of policy on telemedicine utilization [35]. By systematically reviewing these ecosystem roles, this study addresses gaps in previous research that often overlooked the interconnectedness of these factors, providing a more comprehensive understanding of how they collectively drive success in telemedicine healthcare businesses.

The proposed 4Ps fishbone conceptual framework elucidates the critical elements and roles of the entrepreneurial ecosystem that drive success in telemedicine healthcare businesses, encompassing formal institutions, entrepreneurial culture, demand, physical infrastructure, networks, talent, leadership, finance, new knowledge, intermediary services, and productive entrepreneurship. This framework aligns with findings from previous studies, such as those by Kim et al., who emphasize the necessity of a supportive infrastructure and the role of community demand in enhancing telemedicine services [47]. Similarly, Ghiwaa's systematic review highlights the importance of organizational factors and stakeholder engagement in successfully adopting telemedicine [48]. Furthermore, Těšinová's research underscores the significance of user trust and the need for a conducive policy environment, which resonates with the identified roles of process optimization and policy framework in the current study [46]. By integrating these elements into a cohesive framework, this research addresses gaps in earlier studies that

often isolated these factors, thus providing a holistic view of how the entrepreneurial ecosystem can effectively foster success in telemedicine healthcare businesses.

The research presented in this study has certain limitations that should be acknowledged. Firstly, the findings are primarily based on literature sourced exclusively from Scopus, which may limit the scope of evidence by excluding potentially relevant studies from other databases. This constraint could affect the comprehensiveness of the review, particularly in capturing diverse perspectives outside of Scopus-indexed publications. Additionally, the geographical focus of many included studies may impact the generalizability of the results, restricting their application to a broader global context.

This review enriches the knowledge by illustrating how the entrepreneurial ecosystem drives success in telemedicine healthcare businesses by identifying key elements such as institutions, networks, and leadership. The findings offer practical managerial and entrepreneurial implications, enabling healthcare entrepreneurs to leverage these elements for fostering innovation, process optimization, and growth. Furthermore, the 4Ps fishbone conceptual framework is a strategic guide for policymakers to create supportive policies that facilitate ecosystem advancement, ensuring robust development in telemedicine and entrepreneurship within healthcare. This research ultimately benefits society by promoting scalable and efficient telemedicine solutions, enhancing access to quality healthcare services.

5 Conclusion

This study highlights the essential main elements and roles of the entrepreneurial ecosystem in supporting the success of telemedicine businesses. The main elements of the entrepreneurial ecosystem, such as formal institutions, demand, entrepreneurial culture, networks, leadership, talent, finance, physical infrastructure, intermediary services, new knowledge, and productive entrepreneurship, are the foundation for driving innovation and ensuring the scalability of telemedicine services. In this context, the prominent roles of the entrepreneurial ecosystem in the telemedicine health service business include people development, process optimization, policy framework, and platform advancement. Therefore, this study proposes and visualizes the 4Ps fishbone conceptual framework as the principal elements and roles of the entrepreneurial ecosystem that drive success in telemedicine healthcare businesses. Future research should explore the micro-level dynamics of how individual entrepreneurial traits and capabilities within telemedicine businesses interact with ecosystem elements to drive success. At the meso level, studies could investigate the role of organizational structures and networks in optimizing the integration of telemedicine services within healthcare ecosystems. On a macro scale, examining how national policies, regulatory frameworks, and economic environments influence the broader entrepreneurial ecosystem will provide insights into creating supportive conditions for sustainable entrepreneurship in healthcare. Comparative cross-country analyses could further illuminate how varied ecosystem configurations impact the success of telemedicine initiatives. Such studies would enrich the systematic review's findings, guiding policymakers and stakeholders in tailoring strategies for healthcare entrepreneurship.

Acknowledgments. This manuscript is a publication output from an undergraduate thesis track on scientific articles in reputable international publications at the Entrepreneurship Department, Bina Nusantara University, Malang Campus. The author would like to thank the Entrepreneurship Department and Research and Technology Transfer Office, Bina Nusantara University, for supporting the research ecosystem, facilities, and publication funding.

Author Contributorship. **Puan Ayu Maharani**: Data Curation, Formal Analysis, Investigation, Methodology, Resources, Software, Visualization, Writing–Original Draft Preparation, Writing–Review and Editing; **Mardhatillah Shanti**: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Resources, Validation, Funding Acquisition, Project Administration, Supervision, Writing–Review and Editing; **Agung Purnomo**: Conceptualization, Formal Analysis, Methodology, Resources, Visualization, Validation, Project Administration, Supervision, Writing–Review and Editing.

Data Availability. **Harvard Dataverse**: Dataset of Entrepreneurial Ecosystem Drives Success in Telemedicine Healthcare Businesses. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/5ZQLAO (Maharani, P.A., Shanti, M., Purnomo, A., 2024).

References

1. Giunti G, Giunta DH, Guisado-Fernandez E, Bender JL, Fernandez-Luque L (2018) A biopsy of breast cancer Mobile applications: state of the practice review. Int J Med Inform 110:1–9. https://doi.org/10.1016/j.ijmedinf.2017.10.022
2. Lankarani KB et al (2023) The revolutionary impacts of the COVID-19 pandemic on Mobile health; highlights of the fourth shiraz international congress on Mobile health. J Biomed Phys Eng 13:203–208. https://doi.org/10.31661/jbpe.v0i0.2107-1374
3. Moore AF, Hawarden V (2020) Discovery digital health strategy: COVID-19 accelerates online health Care in South Africa. Emerald Emerg Mark Case Stud 10:1–18. https://doi.org/10.1108/EEMCS-06-2020-0197
4. Mandil A, Mabry R, Milani B, Nour M, Afifi M, Abdel-Ghani K (2022) Mapping of health innovations in response to the COVID-19 pandemic in eastern Mediterranean and selected Arab countries. East Mediterr Heal J 28:130–143. https://doi.org/10.26719/emhj.22.028
5. Purnomo A, Afia N, Rosyidah E, Maulana FI, Sugiantoro ADS (2023) Mapping of E-entrepreneurship research: bibliometric review. E3S Web Conf 426:02013. https://doi.org/10.1051/e3sconf/202342602013
6. Lokman L, Chahine T (2021) Business models for primary health care delivery in low- and middle-income countries: a scoping study of nine social entrepreneurs. BMC Health Serv Res 21. https://doi.org/10.1186/s12913-021-06225-6
7. Marzuki N, Ismail S, Al-Sadat N, Ehsan FZ, Chan CK, Ng CW (2015) Integrating information and communication Technology for Health Information System Strengthening: a policy analysis. Asia-Pacific J Public Health 27:865–93S. https://doi.org/10.1177/1010539515590180
8. Chakraborty I, Ilavarasan PV, Edirippulige S (2023) Critical success factors of Startups in the E-health domain. Health Policy Technol 12:100773. https://doi.org/10.1016/j.hlpt.2023.100773
9. Barutçu S (2019) M-health apps design using quality function deployment. Int J Health Care Qual Assur 32:698–708. https://doi.org/10.1108/IJHCQA-08-2018-0195
10. Kok MO et al (2023) Using telehealth to support community health Workers in Uganda during COVID-19: a mixed-method study. BMC Health Serv Res 23. https://doi.org/10.1186/s12913-023-09217-w

11. Jin Y, Ji S (2018) Mapping hotspots and emerging trends of business model innovation under networking in internet of things. EURASIP J Wirel Commun Netw 2018. https://doi.org/10.1186/s13638-018-1115-4
12. Hermes S, Riasanow T, Clemons EK, Böhm M, Krcmar H (2020) The digital transformation of the healthcare industry: exploring the rise of emerging platform ecosystems and their influence on the role of patients. Bus Res 13:1033–1069. https://doi.org/10.1007/s40685-020-00125-x
13. Zilbermint M (2021) The endocrine hospitalist: enhancing the quality of diabetes care. J Diabetes Sci Technol 15:762–767. https://doi.org/10.1177/19322968211007908
14. Alzagladi B (2022) Factors influencing readiness telemedicine implementation during COVID-19: a systematic review. J Kedokt Yars 29:067–073. https://doi.org/10.33476/jky.v29i2.1895
15. Mohammadzadeh N, Rezayi S, Saeedi S (2023) Telemedicine for patient Management in Remote Areas and Underserved Populations. Disaster Med Public Health Prep 17:e167. https://doi.org/10.1017/dmp.2022.76
16. Waseem N, Boulanger M, Yanek LR, Feliciano JL (2022) Disparities in telemedicine success and their association with adverse outcomes in patients with thoracic cancer during the COVID-19 pandemic. JAMA Netw Open 5:e2220543. https://doi.org/10.1001/jamanetworkopen.2022.20543
17. Mabeza RMS, Maynard K, Tarn DM (2022) Influence of synchronous primary care telemedicine versus in-person visits on diabetes, hypertension, and hyperlipidemia outcomes: a systematic review. BMC Prim Care 23:52. https://doi.org/10.1186/s12875-022-01662-6
18. Kia AS, Rafizadeh M, Shahmoradi L (2023) Telemedicine in the emergency department: an overview of systematic reviews. J Public Health (Bangkok) 31:1193–1207. https://doi.org/10.1007/s10389-021-01684-x
19. 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
20. Magnani G, Gioia D (2023) Using the Gioia methodology in international business and entrepreneurship research. Int Bus Rev 32:102097. https://doi.org/10.1016/j.ibusrev.2022.102097
21. Page MJ et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. n71 doi:https://doi.org/10.1136/bmj.n71
22. Aryanto R, Karmagatri M, Purnomo A, Kuncoro EA (2023) Marketing in wireless communication: a systematic review. In: 3rd IEEE international conference on ict in business industry and government, ICTBIG 2023. IEEE, pp 1–5. doi:https://doi.org/10.1109/ICTBIG59752.2023.10456354.
23. Aryani N et al (2024) A systematic review of disruptive innovation on entrepreneurship. Procedia Comput Sci 234:1236–1243. https://doi.org/10.1016/j.procs.2024.03.120
24. Permana AV, Purnomo A, Sarjono H, Maulana FI, Setyani EA (2023) The utilization of Mobile communication on marketing: a systematic review. Procedia Comput Sci. pp 101–109. Elsevier, Malang. doi:https://doi.org/10.1016/j.procs.2023.10.507.
25. Sharma AA, Lee KC, Garibyan L (2022) A call to action: why medical education curriculum needs to encourage Young physicians to innovate. Arch Dermatol Res 314:499–501. https://doi.org/10.1007/s00403-020-02180-0
26. Malakhov KS (2023) Insight into the digital health system of Ukraine (eHealth): trends, definitions, standards, and legislative revisions. Int J Telerehabilitation 15. https://doi.org/10.5195/ijt.2023.6599
27. Salter B, Dong Y, Hunter BM (2022) Constructing healthcare services markets: networks, brokers and the China-England engagement. Glob Health 18. https://doi.org/10.1186/s12992-022-00892-8

28. Wamuyu PK (2015) The impact of information and communication technology adoption and diffusion on technology entrepreneurship in developing countries: the case of Kenya. Inf Technol Dev 21:253–280. https://doi.org/10.1080/02681102.2014.948372
29. Elbanna A, Linderoth HCJ (2015) The formation of technology mental models: the case of voluntary use of Technology in Organizational Setting. Inf Syst Front 17:95–108. https://doi.org/10.1007/s10796-014-9513-6
30. Beckett RC (2017) Hearing enhancement technologies and issues of self-management: a case study. Int J Technol Mark 12:252. https://doi.org/10.1504/IJTMKT.2017.089648
31. de los Reyes G (2019) Institutional entrepreneurship for digital public health promotion: challenges and opportunities. Heal Educ Behav 46:30–36. https://doi.org/10.1177/1090198119871245
32. Velayati F, Ayatollahi H, Hemmat M, Dehghan R (2021) Key components and critical factors for developing a telehealth business framework: a qualitative study. BMC Med Inform Decis Mak 21:339. https://doi.org/10.1186/s12911-021-01707-3
33. Sundin P, Callan J, Mehta K (2016) Why do entrepreneurial MHealth ventures in the developing world fail to scale? J Med Eng Technol 40:444–457. https://doi.org/10.1080/03091902.2016.1213901
34. Parthasarathy M, Khuntia J, Stacey R (2021) Impact of remote and virtual care models on the sustainability of small health care businesses: perceptual analysis of small clinics, physician offices, and pharmacies in Colorado. J Med Internet Res 23. https://doi.org/10.2196/23658
35. Alami H, Gagnon MP, Wootton R, Fortin JP, Zanaboni P (2017) Exploring factors associated with the uneven utilization of telemedicine in Norway: a mixed methods study. BMC Med Inform Decis Mak 17:180. https://doi.org/10.1186/s12911-017-0576-4
36. Cegarra-Navarro J-G, Sánchez-Polo MT (2010) Implementing telemedicine through E-listening in hospital-in-the-home units. Int J Inf Manag 30:552–558. https://doi.org/10.1016/j.ijinfomgt.2010.04.005
37. Stier-Jarmer M et al (2020) A multimodal stress-prevention program supplemented by telephone-coaching sessions to reduce perceived stress among German farmers: results from a randomized controlled trial. Int J Environ Res Public Health 17:1–18. https://doi.org/10.3390/ijerph17249227
38. 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
39. Lam JA, Dang LT, Phan NT, Trinh HT, Vu NC, Nguyen CK (2018) Mobile health initiatives in Vietnam: scoping study. JMIR Mhealth Uhealth 6. https://doi.org/10.2196/mhealth.8639
40. Ashtarian K, Etemadi M (2023) Popular diffusion as an instrument for overcoming barriers to digital health in Iran: the critical role of the pandemic. Int J Health Gov 28:249–266. https://doi.org/10.1108/IJHG-10-2022-0094
41. Gunartin Pratikto H, Winarno A, Restuningdiah N (2023) The role of entrepreneurial competencies: successful key Smes a literature review. Int J Prof Bus Rev 8:e01955. https://doi.org/10.26668/businessreview/2023.v8i7.1955
42. Nti NK (2022) Entrepreneurial management as a predictor of business performance in the Greater Accra region, Ghana. J Manag Entrep Res 3:1–13. https://doi.org/10.34001/jmer.2022.6.03.1-24
43. Wahyuningtyas SY, Hanoteau J, Vial V (2018) The role of social networks in entrepreneurial ecosystems: the case of #StartupLokal in Jakarta. Glob Bus Organ Excell 37:22–29. https://doi.org/10.1002/joe.21889
44. Althumairi A, AlHabib AF, Alumran A, Alakrawi Z (2022) Healthcare providers' satisfaction with implementation of telemedicine in ambulatory care during COVID-19. Healthcare 10:1169. https://doi.org/10.3390/healthcare10071169

45. Wernhart A, Gahbauer S, Haluza D (2019) eHealth and telemedicine: practices and beliefs among healthcare professionals and medical students at a medical university. PLoS One 14:e0213067. https://doi.org/10.1371/journal.pone.0213067
46. Těšinová JK, Dobiášová K, Dušek Z, Tobiášová A (2023) Development of telemedicine in the Czech Republic from patients' and other key stakeholders' perspective. Front Public Health 11. https://doi.org/10.3389/fpubh.2023.1202182
47. Kim HY et al (2018) Health service utilisation, unmet healthcare needs, and the potential of telemedicine services among Korean expatriates. Glob Health 14:120. https://doi.org/10.1186/s12992-018-0433-y
48. Ghiwaa T, Khan I, White M, Beloff N (2023) Telemedicine adoption for healthcare delivery: a systematic review. Int J Adv Comput Sci Appl 14. https://doi.org/10.14569/IJACSA.2023.01411125

Leveraging Medical Informatics for Data-Driven Entrepreneurial Ventures in the Healthcare Sector: A Systematic Review

Wafiq Azizah, Riefky Prabowo(✉), and Agung Purnomo

BINUS Business School Undergraduate Program, Entrepreneurship Department, Bina Nusantara University, Jakarta 11480, Indonesia
riefky.prabowo@binus.ac.id

Abstract. The rapid integration of data-driven healthcare services with medical informatics is reshaping entrepreneurship in the healthcare sector. Yet, there is limited understanding of the internal and external factors that drive success in this area. This study aims to identify the development of global literature on internal and external factors that are key to the success of implementing medical information for data-driven entrepreneurship in the healthcare business. Unlike previous studies, this research systematically integrates internal and external factors to establish a comprehensive framework for data-driven success in healthcare entrepreneurship through medical informatics. Using the Systematic Literature Review (SLR) method, peer-reviewed papers published were analyzed from the Scopus database, focusing on medical informatics and entrepreneurship. The findings reveal four key internal success factors: data-driven capabilities, digital transformation leadership, a dynamic collaborative ecosystem, and a driven entrepreneurial spirit. Four external factors were identified: directive and resource environment, competitive dynamics, policy support, and digital training integration. Based on the research findings, this study proposes the 8D padlock conceptual framework as the key success factor of medical informatics for data-driven entrepreneurship in the health business. These results highlight the importance of enhancing data-driven capabilities, supporting digital transformation, fostering collaboration, and ensuring continuous investment in digital training. Policymakers and entrepreneurs can use these insights to create more sustainable and innovative healthcare ventures.

Keywords: Entrepreneur · Entrepreneurship · Healthcare · Medical Informatics · Systematic Literature Review

1 Introduction

Business-based medical informatics has grown significantly, driven by technological advancements and the demand for data-driven healthcare. The Internet of Medical Things is expanding through the convergence of medical equipment, infrastructure, and the Internet, improving patient data management, healthcare efficiency, and clinical outcomes [1, 2]. Integrating health information management and data analytics into advanced

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 27–39, 2026.
https://doi.org/10.1007/978-981-95-9233-3_3

academic programs highlights the evolving competencies required in this field [3]. Substantial investments in electronic health records (EHRs), driven by healthcare legislative acts, further emphasize the need for entrepreneurial skills in this growing domain [4].

Healthcare entrepreneurship is expanding, with a growing focus on the role of clinicians in driving innovation. Clinician engagement in entrepreneurship has been shown to foster more effective and financially viable healthcare solutions [5]. Medical informatics programs enhance clinicians' innovative capacities, paving the way for healthcare startups, while strategic positioning of personal health records (PHRs) creates new opportunities for entrepreneurship in healthcare [6, 7].

The urgency of studying medical informatics for data-driven entrepreneurship lies in leveraging data for better healthcare outcomes. Advances in AI and machine learning technologies could present opportunities to improve clinical decision-making, necessitating an entrepreneurial mindset among medical professionals [8]. Overcoming challenges in integrating AI into clinical practice, including regulatory and organizational barriers, is crucial for successfully adopting these technologies [9]. The SLR method is ideal for researching entrepreneurship ecosystems in medical informatics, providing a comprehensive approach to methodically review the available literature, ascertain essential issues, and further investigate the interplay between technological advancements and entrepreneurial activities [10, 11]. This method can reveal critical insights into the factors driving innovation and the barriers within healthcare ecosystems.

Previous research has used various methods to study internal and external factors influencing data-driven entrepreneurial success in medical informatics. Unertl et al. utilized a comparative study, which, while effective in identifying contextual challenges, lacked depth in exploring the nuances of informatics integration [12]. Alhosseini et al. conducted a content analysis highlighting barriers in Iranian medical universities; however, its limited geographical focus restricts generalizability [13]. Ismail et al. employed qualitative methods to assess pharmacists' roles, yet the study's reliance on a single professional perspective may overlook broader systemic factors [14]. Luna et al. reviewed health informatics challenges in developing countries, but the lack of empirical data weakens its applicability [15]. Borycki emphasized the need for an evidence base in eHealth but did not systematically analyze the interplay of various factors [16]. Velayati et al. developed a telehealth business framework through qualitative research; however, the absence of quantitative validation raises questions about its robustness [6]. These studies collectively highlight a research gap in systematically analyzing the interplay of internal and external factors affecting entrepreneurial success in medical informatics.

The systematic literature review employing grounded theory and thematic analysis based on the Gioia method can address these weaknesses by integrating diverse methodologies and providing a comprehensive framework that captures the complexities of the field [17]. This approach enhances theoretical insights and practical implications, distinguishing it from prior research by offering a more holistic understanding of the factors driving entrepreneurial success in medical informatics. Therefore, this study aims to identify the development of global literature on internal and external factors that are key to the success of implementing medical information for data-driven entrepreneurship in the healthcare business. This significant study identifies critical internal and external success factors for leveraging medical informatics in data-driven entrepreneurship

within the healthcare sector. The insights gained from this research will support healthcare practitioners, entrepreneurs, and policymakers in fostering sustainable, innovative ventures that ultimately benefit society by improving healthcare accessibility, efficiency, and quality.

2 Materials and Method

A Systematic Literature Review (SLR) was conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) guidelines in this medical information for data-driven entrepreneurship in healthcare study. The PRISMA framework was selected because it offers a rigorous, structured approach, ensuring the review is transparent and systematic [18]. This protocol also includes specific criteria for selecting, determining eligibility, and excluding articles, which is essential for thoroughly covering the relevant literature and effectively synthesizing evidence [19]. Data for this study was gathered from the Scopus Database, known for its extensive and high-quality coverage, particularly in chemical engineering and entrepreneurship, making it highly relevant for this review [20].

The search command keywords in the document search have been "medical informatics" and "entrepreneurship," combined with the corresponding terms. Twenty-one documents were reviewed and met the inclusion criteria for medical informatics and entrepreneurship and publications in the article, as seen in Table 1. The search query option used in data mining was as follows (TITLE-ABS-KEY ("medical Informatic*" OR "health informatic*" OR "clinical informatic*") AND TITLE-ABS-KEY (entrepreneur*)) AND (LIMIT-TO (DOCTYPE, "ar")). The dataset of this study can be found at the end of the manuscript.

Table 1. Inclusion criteria

Inclusion criteria	Description
Topics	Address medical informatics and entrepreneurship
Document types	Article

Table 1 outlines inclusion criteria limited to studies addressing "medical informatics and entrepreneurship" and document types restricted to "articles." The authors chose not to limit by publication year because doing so would have significantly reduced the pool of relevant studies, yielding fewer than 100 papers from the Scopus database. The research's reliance on these two criteria—topic relevance and document type—ensures a focused exploration of literature that directly contributes to understanding the role of medical informatics in entrepreneurship while maintaining a sufficient dataset for a robust systematic review [18]. The PRISMA protocol flow diagram of medical information implementation for data-driven entrepreneurship in healthcare studies can be seen in Fig. 1.

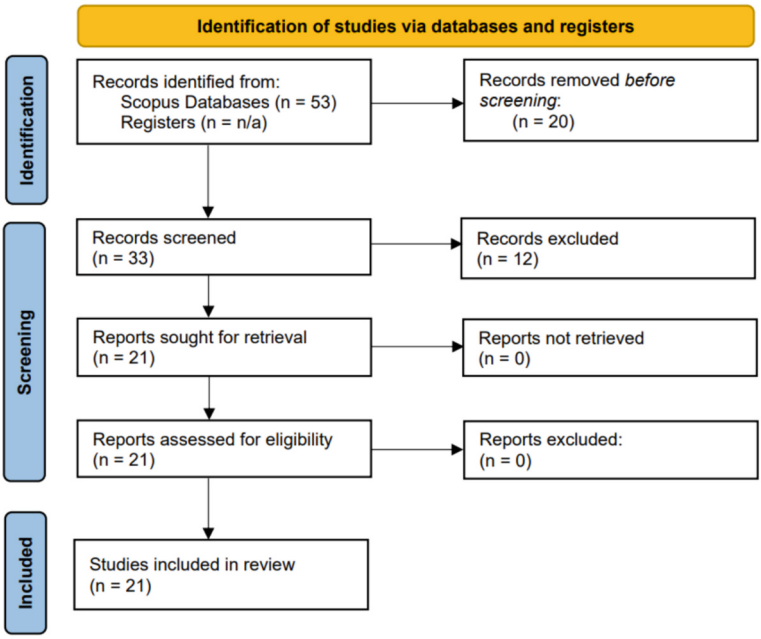


Fig. 1. PRISMA protocols

The research questions (RQs) guiding the medical information implementation for data-driven entrepreneurship in healthcare study were structured as follows.

- RQ1: What internal factors are the key to success in implementing medical information for data-driven entrepreneurship in the healthcare business?
- RQ2: What external factors are the key to success in implementing medical information for data-driven entrepreneurship in the healthcare business?
- RQ3: What are the future research opportunities of medical informatics for entrepreneurial ventures in the healthcare business from micro, meso, and macro perspectives?

In the analysis, RQ1 and RQ2 were examined using descriptive and thematic approaches, employing the Gioia method. This method offers a structured way to analyze qualitative data, which helps to gain a more profound understanding of the internal and external factors that are key to the success of medical information for data-driven entrepreneurship. The Gioia method provides a systematic framework for organizing and interpreting textual data, making it easier to extract meaningful insights from the literature [17]. Furthermore, RQ3 extends this analysis by identifying future research opportunities in leveraging medical informatics for entrepreneurial ventures in healthcare, allowing for a forward-looking perspective on the evolution of this field.

3 Results

3.1 The Internal Key Success Factors

The thematic structuring and analysis process using the Gioia method identified four aggregate dimensions that identify internal key success factors of implementing medical information for data-driven entrepreneurship in the healthcare business, as shown in Fig. 2. These dimensions include data-driven capabilities, digital transformation leadership, a dynamic collaborative ecosystem, and a driven entrepreneurial spirit. Each dimension critically shapes the success of entrepreneurship in the medical information field. Detailed explanations of these dimensions are provided in the following sections.

Data-Driven Capabilities.
The Data-Driven Capabilities dimension is key to medical information entrepreneurship. It focuses on effective health data management, analytics, and AI skills to improve decision-making [3, 10]. Genomic and informatics integration enhances research and market applications [21, 22]. Evidence-based practices and knowledge management drive informed decisions [21, 23]. Public health, policy expertise, and translational research ensure solutions meet standards [7, 9]. These internal factors boost innovation and success in data-driven entrepreneurship.

Digital Transformation Leadership.
The dimensions of digital transformation leadership emphasize solid technological skills, informatics expertise, and effective leadership to drive digital change [3, 24, 25]. Continuous learning and strategic management are crucial to adapting to a fast-paced environment [24, 26]. Change management and systems thinking guide organizational transformation, while collaboration and user-centered design foster innovation and engagement [23, 25]. These internal factors collectively support successful digital transformation in medical entrepreneurship.

Dynamic Collaborative Ecosystem.
The Dynamic Collaborative Ecosystem dimension centers on collaboration across disciplines to drive innovation in medical entrepreneurship [3, 10]. Cooperation between academia and industry is vital, along with user-centered design and strong communication for stakeholder engagement [9, 24, 25]. Strategic workforce management and relationship building to support operational efficiency [27]. These internal factors create an environment that fosters continuous innovation and success in data-driven medical entrepreneurship.

Driven Entrepreneurial Spirit.
Driven entrepreneurial spirit in digital health entrepreneurship is essential [3]. Innovation in health technologies drives progress and keeps businesses competitive [22]. Passion for change and a robust entrepreneurial attitude are necessary for adapting to new challenges [26]. Self-efficacy beliefs and perceived behavioral control empower entrepreneurs to take decisive actions [28]. Developing an entrepreneurial mindset builds resilience and strategic thinking [29]. These factors collectively drive vision, enabling entrepreneurs to excel in medical information and data-driven entrepreneurship, staying ahead in a fast-evolving field.

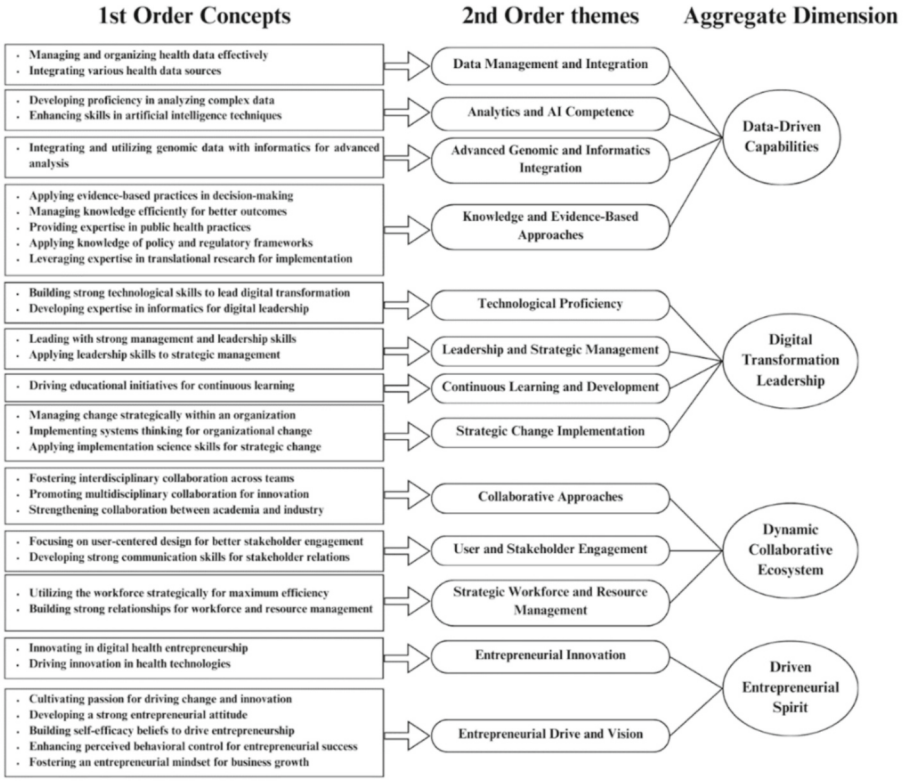


Fig. 2. The internal key success factors of medical informatics for data-driven entrepreneurship

3.2 The External Key Success Factors

The thematic structuring and data analysis using the Gioia method identified four aggregate dimensions representing the external key success factors of implementing medical information for data-driven entrepreneurship in the healthcare business, as shown in Fig. 3. These dimensions were directive and resource environment, dynamic and competitive dynamics, development and policy support, and digital and training integration. Each dimension addresses crucial aspects such as regulatory compliance, resource access, fostering collaboration, maintaining competitive advantage, supportive policies, and technological integration. The following sections provide a detailed explanation of each aggregate dimension.

Directive and Resource Environment.
The Directive and Resource Environment are essential for the success of data-driven medical entrepreneurship. Regulatory frameworks ensure compliance and foster innovation, while ethical guidelines protect patient data [3, 11]. Financial resources, including venture capital, support startups and research [3, 4]. Academic institutions and educational collaborations drive innovation in health informatics [3, 5]. Access to funding and resource partnerships strengthen the ecosystem. A supportive academic environment

further advances research and development efforts. Directive and resource availability are critical factors in driving the growth and adoption of medical information ventures.

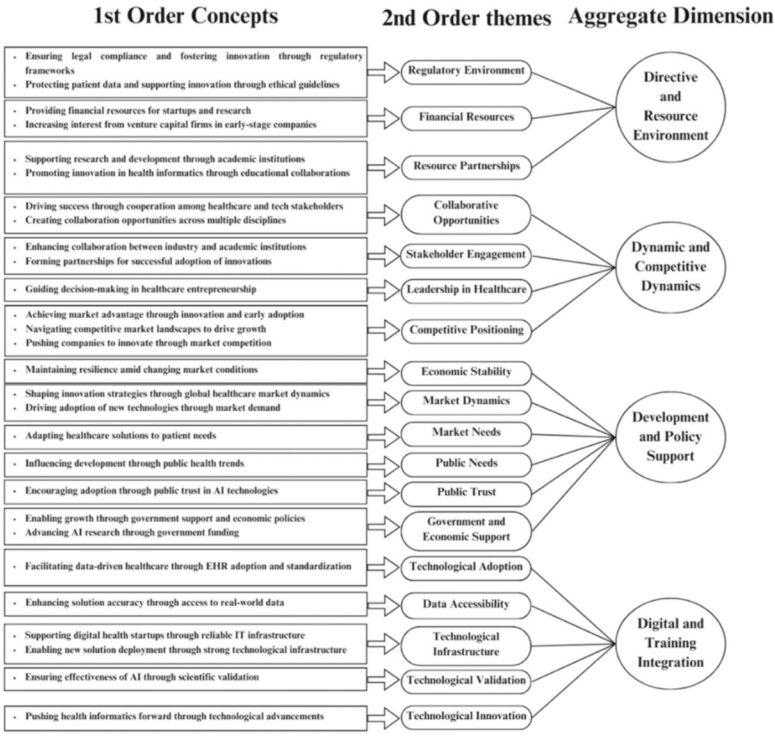


Fig. 3. The external key success factors of medical informatics for data-driven entrepreneurship

Dynamic and Competitive Dynamics.
Dynamic and Competitive Dynamics play a vital role in the success of data-driven medical entrepreneurship. Collaboration with stakeholders and interdisciplinary opportunities drive innovation [26, 30]. Partnerships with industry and academic institutions enhance development [27, 28]. Guiding decision-making in healthcare entrepreneurship helps achieve market advantage through innovation and early adoption [9, 25]. Navigating competitive landscapes and market competition push continuous improvement [7, 31]. Stakeholder engagement is crucial for implementing new solutions. Overall, dynamic and competitive factors fuel success in the healthcare sector.

Development and Policy Support.
The Development and Policy Support dimension is essential for data-driven medical entrepreneurship. Economic resilience enables adaptation to market changes [30]. Global healthcare dynamics influence innovation strategies [21]. Market demand drives technology adoption, while understanding patient needs tailors solutions [23, 27]. Public

health trends shape development, and public trust in AI is vital for adoption [8, 26]. Government support, funding, and policies lay the foundation for growth and innovation in healthcare [6, 10].

Digital and Training Integration.
The Digital and Training Integration dimension is crucial for data-driven medical entrepreneurship. Standardized EHR adoption supports data-driven healthcare [4]. Access to real-world data enhances solution accuracy [29]. Reliable IT infrastructure and solid technological systems enable growth and implementation [24, 32]. Scientific validation ensures AI effectiveness in healthcare [8]. Continuous technological advancements drive innovation [22]. Digital and proper training are essential to successfully integrating new technologies in the healthcare sector.

3.3 Key Findings

Based on the research findings, this study identifies eight critical success factors of medical informatics for data-driven entrepreneurship in health business. Internally, success is driven by four factors: data-driven capabilities, digital transformation leadership, a dynamic collaborative ecosystem, and a driven entrepreneurial spirit. Externally, success depends on four factors: directive and resource environment, development and policy support, adapting through dynamic and competitive dynamics, and digital and training integration. Therefore, this study proposes and visualizes the 8D padlock conceptual framework as the critical success factor of medical informatics for data-driven entrepreneurship in the health business, as shown in Fig. 4. These factors collectively ensure growth, innovation, and resilience in the entrepreneurial venture of the healthcare business.

4 Discussion

Analyzing internal factors critical for implementing medical information in data-driven entrepreneurship within the healthcare sector reveals four aggregate dimensions: data-driven capabilities, digital transformation leadership, a dynamic collaborative ecosystem, and a driven entrepreneurial spirit. These findings align with previous research emphasizing the importance of organizational resources and capabilities in fostering innovation and entrepreneurship [33]. For instance, Menear et al. highlights the necessity of adaptable leadership and collaborative frameworks in learning health systems, which resonates with our identified dimensions of leadership and collaboration [34]. Furthermore, the emphasis on data-driven capabilities is echoed in the work of Øvretveit et al., who discuss the transformative potential of clinical data in enhancing healthcare delivery [35]. Unlike earlier studies that primarily focused on external factors or broad organizational resources, our systematic review provides a nuanced understanding of specific internal factors that directly influence the success of medical informatics initiatives, thereby addressing gaps in the literature regarding the operationalization of these concepts in entrepreneurial contexts [36].

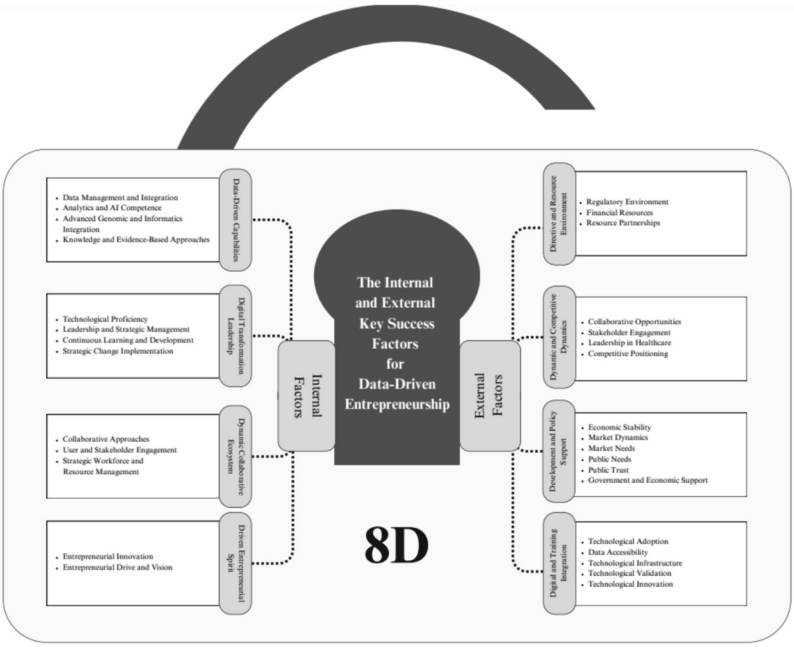


Fig. 4. The 8D padlock conceptual framework of key success factors

The analysis of external factors influencing the successful implementation of medical information for data-driven entrepreneurship in healthcare has identified four critical dimensions: directive and resource environment, dynamic and competitive dynamics, development, policy support, and digital and training integration. These findings are consistent with previous studies that emphasize the role of supportive institutional frameworks and resource availability in fostering entrepreneurial activities. For instance, Nasiri and Hamelin highlight the significance of educational contexts in shaping entrepreneurial opportunities, particularly in the MENA region [37]. Similarly, Haldar and Tripathi discuss how external uncertainties, driven by institutional and technological factors, can significantly impact entrepreneurship in the renewable energy sector [38]. Furthermore, the emphasis on policy support aligns with findings from Mishra and Pandey, who underscore the necessity of conducive regulatory environments for healthcare entrepreneurship [39]. Unlike earlier studies that often focused on isolated external factors, our systematic review integrates these dimensions into a cohesive framework, addressing the multifaceted nature of external influences on data-driven entrepreneurship in healthcare, thus filling a critical gap in the literature.

The proposed 8D padlock conceptual framework elucidates the critical success factors for leveraging medical informatics in data-driven entrepreneurship within the healthcare sector. Internally, the framework identifies data-driven capabilities, digital transformation leadership, a dynamic collaborative ecosystem, and a driven entrepreneurial spirit as pivotal elements. Externally, it emphasizes the importance of a supportive directive and resource environment, development and policy support, adaptability to dynamic and

competitive dynamics, and integration of digital training. This aligns with findings from Nasiri and Hamelin, who assert that educational attainment enhances opportunity-driven entrepreneurship, suggesting that a well-educated workforce is crucial for fostering innovation in healthcare [37]. Additionally, Haldar and Tripathi highlight the significance of contextual factors, such as institutional support, in shaping entrepreneurial success [38]. Unlike previous studies that often isolated factors or focused on singular aspects of entrepreneurship, our framework integrates both internal and external dimensions, providing a comprehensive approach to understanding the multifaceted nature of success in medical informatics entrepreneurship. This holistic perspective addresses gaps in the existing literature by illustrating how these interconnected factors collectively influence entrepreneurial outcomes in healthcare [40].

Leveraging insights from medical informatics through this systematic review offers substantial contributions to the body of knowledge, particularly by enhancing understanding of the data-driven factors that underpin entrepreneurial ventures in healthcare. This research informs entrepreneurship practices by highlighting the pivotal role of data-driven capabilities and collaborative ecosystems, which support healthcare entrepreneurs in making informed, strategic decisions. For policymakers, it highlights the need for supportive policies, digital training, and resource access to foster healthcare innovation. The study demonstrates that leveraging medical informatics can drive more efficient, accessible healthcare solutions, providing broad societal benefits through improved health outcomes and innovative service delivery.

This systematic review's reliance on the Scopus database may limit the diversity of sources, potentially excluding relevant studies. The focus on medical informatics in healthcare entrepreneurship also restricts generalizability to other sectors. Additionally, the study's qualitative approach lacks empirical validation, which would strengthen the proposed framework. Lastly, emphasis on policy and digital transformation might overlook other entrepreneurial factors, indicating a need for more quantitative research.

5 Conclusion

This systematic literature review identifies key success factors of leveraging medical informatics for data-driven entrepreneurial ventures in the healthcare business. The aim was to explore internal and external factors driving success. Internally, success is driven by data-driven capabilities, digital transformation leadership, a dynamic, collaborative ecosystem, and entrepreneurial spirit. Externally, it relies on a directive and resource environment, dynamic competition, development and policy support, and digital training integration. This study proposes and visualizes the 8D padlock conceptual framework as a key success factor of medical informatics for data-driven entrepreneurship in the health business. Future research should focus on developing scalable and sustainable models for integrating medical informatics into healthcare entrepreneurship. From a micro perspective, standardized frameworks and ethical approaches to patient data privacy are essential to ensure proper and secure adoption. At the meso level, collaboration between academia, industry, and healthcare institutions is crucial for translating research into practical, innovative applications that align with market needs. From a macro perspective, government policy support and investment in digital transformation,

particularly in artificial intelligence and predictive analytics, are vital to drive innovation and sustainability in data-driven entrepreneurship within the healthcare sector.

Acknowledgments. This manuscript is a publication output from an undergraduate thesis track on scientific articles in reputable international publications at the Entrepreneurship Department, Bina Nusantara University, Malang Campus. The author would like to thank the Entrepreneurship Department and Research and Technology Transfer Office, Bina Nusantara University, for supporting the research ecosystem, facilities, and publication funding.

Author Contributorship. Wafiq Azizah: Data Curation, Formal Analysis, Investigation, Methodology, Resources, Software, Visualization, Writing–Original Draft Preparation, Writing–Review and Editing; **Riefky Prabowo**: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Resources, Validation, Funding Acquisition, Project Administration, Supervision, Writing–Review and Editing; **Agung Purnomo**: Conceptualization, Formal Analysis, Methodology, Resources, Visualization, Validation, Project Administration, Supervision, Writing–Review and Editing.

Data Availability. Harvard Dataverse: Dataset of Medical Informatics for Data-Driven Entrepreneurial Ventures in the Healthcare. https://doi.org/10.7910/DVN/EACJA0 (Azizah, W., Prabowo, R., Purnomo, A., 2024).

References

1. Maulana FI, Lestari D, Adi PDP, Arifin MN, Purnomo A (2024) Internet of medical things: a bibliometric analysis of research publications from 2018–2022. In: Lecture notes in electrical engineering. Springer, Singapore, pp 395–408. https://doi.org/10.1007/978-981-97-1463-6_27
2. Purnomo A, Firdaus M, Maulana FI, Triangga B, Hikmawan MI, Hikmah ZTN (2024) Business intelligence in healthcare: a review of knowledge structures and level of analysis. In: Lecture notes in electrical engineering. Springer, Singapore, pp 505–517. https://doi.org/10.1007/978-981-97-1463-6_35
3. Ahmadi M, Sheikhtaheri A, Tahmasbi F, Eslami Jahromi M, Rangraz Jeddi F (2022) A competency framework for Ph.D. programs in health information management. Int J Med Inform 168:104906. https://doi.org/10.1016/j.ijmedinf.2022.104906
4. Lite S, Gordon WJ, Stern AD (2020) Association of the Meaningful use Electronic Health Record Incentive Program with Health Information Technology Venture Capital Funding. JAMA Netw Open 3:e201402. https://doi.org/10.1001/jamanetworkopen.2020.1402
5. Archer A, Mcneil J, Johnson T, Ferlie E, Nagy P (2022) Impact of entrepreneurship training on clinician engagement in innovation creation: an evaluation of the Johns Hopkins Hexcite programme. BMJ Lead 6:50–52. https://doi.org/10.1136/leader-2019-000197
6. Velayati F, Ayatollahi H, Hemmat M, Dehghan R (2021) Key components and critical factors for developing a telehealth business framework: a qualitative study. BMC Med Inform Decis Mak 21:339. https://doi.org/10.1186/s12911-021-01707-3
7. Lewis M, Baxter R, Pouder R (2013) The development and deployment of electronic personal health records. J Health Organ Manag 27:577–600. https://doi.org/10.1108/JHOM-07-2012-0144
8. Maassen O et al (2021) Future medical artificial intelligence application requirements and expectations of physicians in German university hospitals: web-based survey. J Med Internet Res 23:e26646. https://doi.org/10.2196/26646

9. Greenblatt WH (2021) Proportion, type, and characteristics of physician entrepreneurship in Massachusetts. JAMA Netw Open 4. https://doi.org/10.1001/jamanetworkopen.2020.26938
10. Liu F, Qu S, Fan Y, Chen F, He B (2023) Scientific creativity and innovation ability and its determinants among medical postgraduate students in Fujian Province of China: a cross sectional study. BMC Med Educ 23:444. https://doi.org/10.1186/s12909-023-04408-9
11. Reiner BI (2008) Intellectual property in medical imaging and informatics: the independent inventor's perspective. J Digit Imaging 21:3–8. https://doi.org/10.1007/s10278-007-9096-6
12. Unertl KM et al (2016) Integrating community-based participatory research and informatics approaches to improve the engagement and health of underserved populations. J American Med Inform Assoc. 23:60–73. https://doi.org/10.1093/jamia/ocv094
13. Sadat Naghavi Alhosseini S., , Pourabbasi, A., Banay Razi, S.: Investigating capacities and barriers of Iranian medical universities in developing entrepreneurship in terms of educational experts: a content analysis study. Med J Islam Repub Iran 35, (2021). https://doi.org/10.47176/mjiri.35.37.
14. Ismail NF, Snoswell C, Banks V, Falconer N (2023) The roles and perspectives of an informatics pharmacist practicing in the Australian healthcare setting: a qualitative study. J Pharm Pract Res 53:116–125. https://doi.org/10.1002/jppr.1860
15. Luna D, Almerares A, Mayan JC, Bernaldo G, de Quirós F, Otero C (2014) Health informatics in developing countries: going beyond pilot practices to sustainable implementations: a review of the current challenges. Healthcare Inform Res 20(3). https://doi.org/10.4258/hir.2014.20.1.3
16. Borycki E (2019) Quality and safety in eHealth: the need to build the evidence base. J Med Internet Res 21:e16689. https://doi.org/10.2196/16689
17. Magnani G, Gioia D (2022) Using the Gioia methodology in international business and entrepreneurship research. Int Bus Rev 31:102097. https://doi.org/10.1016/j.ibusrev.2022.102097
18. Page MJ et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. n71. doi:10.1136/bmj.n71.
19. Aryani N et al (2024) A systematic review of disruptive innovation on entrepreneurship. Procedia Comput Sci 234:1236–1243. https://doi.org/10.1016/j.procs.2024.03.120
20. Permana AV, Purnomo A, Sarjono H, Maulana FI, Setyani EA (2023) The utilization of Mobile communication on marketing: a systematic review. Procedia Comput Sci. Elsevier, Malang, pp 101–109. doi:https://doi.org/10.1016/j.procs.2023.10.507.
21. Epstein RJ (2010) Digitization and its discontents: future shock in predictive oncology. Semin Oncol 37:60–64. https://doi.org/10.1053/j.seminoncol.2009.12.006
22. Poste, G.: Molecular Medicine and Information–Based Targeted Healthcare. Nat Biotechnol 16, 19–21 (1998). doi:https://doi.org/10.1038/5403.
23. Morrato EH et al (2023) The evolve to next-gen ACT network: an evolving open-access, real-world data resource primed for real-world evidence research across the clinical and translational science award consortium. J Clin Transl Sci 7:e224. https://doi.org/10.1017/cts.2023.617
24. Ancker JS, Witteman HO, Hafeez B, Provencher T, Van de Graaf M, Wei E (2015) "you get reminded you're a sick person": personal data tracking and patients with multiple chronic conditions. J Med Internet Res 17:e202. https://doi.org/10.2196/jmir.4209
25. Brimhall BB, Wright LD, Mcgregor KL, Hernandez JS (2007) Critical leadership and management skills for pathology practice. Arch Pathol Lab Med 131:1547–1554
26. Kelley T (2023) Dr. Saba's innovative contributions to nursing informatics. Int J Med Inform 170:104982. https://doi.org/10.1016/j.ijmedinf.2022.104982
27. Schenkat D, Rough S, Hansen A, Chen D, Knoer S (2018) Creating organizational value by leveraging the multihospital pharmacy Enterprise. American J Heal Pharm 75:437–449. https://doi.org/10.2146/ajhp170375

28. Wu W-H, Wei C-W, Yu M-C, Kao H-Y (2020) Exploring factors surrounding students' entrepreneurial intentions in medical informatics: the theory of planning behavior perspective. Front Psychol 11. https://doi.org/10.3389/fpsyg.2020.544887
29. Carroll CA (2004) Enhancing entrepreneurship and professionalism in medical informatics instruction: a collaborative training model. J American Med Inform Assoc 11:317–319. https://doi.org/10.1197/jamia.M1463
30. Kaplan K (2011) California: safe harbour. Nature 472:125–126. https://doi.org/10.1038/nj7341-125a
31. Larvol BL, Wilkerson LJ (1998) In Silico drug discovery: tools for bridging the NCE gap. Nat Biotechnol 16:33–34. https://doi.org/10.1038/5412
32. Ashtarian K, Etemadi M (2023) Popular diffusion as an instrument for overcoming barriers to digital health in Iran: the critical role of the pandemic. Int J Health Gov 28:249–266. https://doi.org/10.1108/IJHG-10-2022-0094
33. Urbano D, Alvarez C, Turró A (2013) Organizational resources and intrapreneurial activities: an international study. Manag Decis 51:854–870. https://doi.org/10.1108/00251741311326617
34. Menear M, Blanchette M-A, Demers-Payette O, Roy D (2019) A framework for value-creating learning health systems. Health Res Policy Syst 17:79. https://doi.org/10.1186/s12961-019-0477-3
35. Ovretveit J, Nelson E, James B (2016) Building a learning health system using clinical registers: a non-technical introduction. J Health Organ Manag 30:1105–1118. https://doi.org/10.1108/JHOM-06-2016-0110
36. Amri MM, Abed SA (2023) The data-driven future of healthcare: a review. Mesopotamian J Big Data 1:70–76. https://doi.org/10.58496/MJBD/2023/010
37. Nasiri N, Hamelin N (2018) Entrepreneurship driven by opportunity and necessity: effects of educations, gender and occupation in MENA. Asian J Bus Res 8. https://doi.org/10.14707/ajbr.180049
38. Haldar S, Tripathi T (2023) What besets entrepreneurs in renewable energy sector? – insights from the Indian state of Gujarat. Int J Energy Sect Manag 17:209–226. https://doi.org/10.1108/IJESM-06-2021-0002
39. Mishra A, Pandey N (2023) Global entrepreneurship in healthcare: a systematic literature review and bibliometric analysis. Glob Bus Organ Excell 42:9–21. https://doi.org/10.1002/joe.22193
40. Wannamakok W, Yonwikai W (2023) The relationship between institutional environments and Thai hospitality-oriented entrepreneurship through the moderating role of educational support: a mixed methods approach. Adm Sci 13:188. https://doi.org/10.3390/admsci13080188

Research Trends in Telemedicine Entrepreneurship: Emerging Mapping from a Global Perspective

Noor Anisa Nabila and Agung Purnomo(✉)

BINUS Business School Undergraduate Program, Entrepreneurship Department, Bina Nusantara University, Jakarta 11480, Indonesia
agung.purnomo@binus.ac.id

Abstract. Telemedicine has proliferated, especially in the digital business and entrepreneurship sector, by delivering remote healthcare services through technology. This study aims to map and analyze the global trends of telemedicine entrepreneurship from 1994 to 2023 using a bibliometric approach. This research contributes by providing a comprehensive bibliometric mapping of telemedicine entrepreneurship, identifying emerging themes and gaps that offer a framework for future innovation and strategic advances in global healthcare. By analyzing 165 documents indexed in Scopus using the R Biblioshiny tool, we explore the thematic map evolution and research trends of telemedicine in entrepreneurship. The results show a substantial increase in scholarly production related to telemedicine entrepreneurship, with an annual growth of 12.18%, reflecting the growing global interest. The three-field analysis plots review the various international contributions, highlighting their interdisciplinary nature and global relevance. Trending topics, motor themes, and word clouds refer to entrepreneurship, innovation, telehealth, telemedicine, healthcare, and digital health research topics. The research implications highlight that understanding global research trends in telemedicine entrepreneurship through emerging thematic mapping offers valuable insights for academia, practical applications, and policymakers, supporting informed decisions and fostering innovation from a global perspective.

Keywords: Bibliometric · Digital Business · Entrepreneurship · Health Technology · Healthcare · Telemedicine

1 Introduction

The development of telemedicine as a healthcare innovation has shown significant progress in recent years, predominantly driven by global challenges such as the COVID-19 pandemic. Telemedicine, which involves medical practices through communication technology, has become an essential solution due to its ability to increase the accessibility of healthcare services, optimize operational efficiency, and reduce treatment costs [1, 2]. By utilizing digital technology, telemedicine allows patients to receive medical consultations without interacting directly with healthcare providers. This innovation has proven very relevant in conditions that limit mobility, such as during a pandemic [3, 4].

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 40–52, 2026.
https://doi.org/10.1007/978-981-95-9233-3_4

The telemedicine business trend focuses on providing remote medical consultation services. It includes various technological supporters, such as developing health applications, wearable devices, and digital platforms that enable more interactive interactions between patients and healthcare providers [5]. Wearable devices, for example, can monitor a patient's health condition in real-time and transmit data directly to medical professionals. This creates vast business opportunities in the healthcare sector, enabling more responsive and proactive healthcare services [6]. This development shows telemedicine has evolved beyond just a diagnostic and treatment tool but also a driver of technological innovation in the healthcare sector [7].

Technology-based entrepreneurship has experienced rapid growth over the past few decades and has positively impacted the healthcare sector. Technology entrepreneurship, including innovative products and disruptive business models, has become essential to global economic growth [8]. These technology-driven business models create new jobs and increase global competitiveness. In telemedicine, technology entrepreneurship offers excellent potential to combine healthcare needs with innovative solutions based on information and communication technologies [9]. This enables entrepreneurs in the healthcare sector to create services and products that are efficient, sustainable, and in line with the evolving needs of modern society [10, 11].

The research problem is the lack of information on global trends and mapping in telemedicine entrepreneurship to meet evolving consumer demands and drive innovation, efficiency, and sustainability in healthcare. The urgency of developing entrepreneurship in telemedicine is increasingly felt along with changes in consumer behavior that demand more flexible and easily accessible healthcare services at all times. Telemedicine offers a practical and efficient solution, especially in situations of limited mobility or traditional healthcare facilities that are difficult to access [12]. Research shows telemedicine can reduce patient waiting times for healthcare services while increasing patient satisfaction. From the perspective of service providers, telemedicine also helps reach more patients without significantly increasing physical service capacity. This improves operational efficiency and provides cost-saving opportunities for patients and service providers [13].

Bibliometric analysis has the potential to be used as the primary research method to explore global trends and maps in telemedicine entrepreneurship [14]. This approach provides a comprehensive quantitative overview of telemedicine entrepreneurship through several vital analyses. Summary statistics offer insight into the overall volume of research, top contributing countries, and lead authors. Annual production performance reveals research growth over time, often reflecting how external events, such as pandemics, have spurred innovation. The three-field plot maps the relationship between authors, institutions, and countries, illustrating the interdisciplinary and international nature of the research. Word cloud analysis visualizes dominant topics such as telehealth, entrepreneurship, and innovation, while trend topic analysis tracks how research themes have evolved. Thematic mapping further identifies mature and emerging research clusters, clarifying where new opportunities lie [15, 16].

Telemedicine entrepreneurship has been explored through various research methods, revealing both advantages and disadvantages across studies. For instance, Sahoo et al. conducted a literature review to assess telemedicine's current state, highlighting its rapid technological advancements but noting a lack of empirical data on adoption barriers [17]. Kozikowski et al. utilized qualitative methods to gather care team perspectives on telemedicine, providing rich insights but limited generalizability due to small sample sizes [18]. Theodore et al. conducted a telemedicine transaction cost analysis, demonstrating cost-effectiveness but not addressing patient satisfaction [19]. Saberi et al. focused on a pilot study, which offered initial insights into telemedicine potential but lacked scalability [20]. Ziegler's scoping review identified telemedicine innovations during the COVID-19 pandemic but did not systematically evaluate their long-term impacts [21]. Jack et al. provided practical telemedicine guidance through qualitative analysis, but their findings were context-specific and may not apply universally [22].

The research gap in telemedicine entrepreneurship lies in the lack of comprehensive bibliometric analyses that synthesize existing literature to identify trends and emerging themes. While qualitative and quantitative studies have provided valuable insights, they often focus on specific contexts or methodologies, leaving a broader understanding of global trends unaddressed. By employing bibliometric methods, this research can systematically map the evolution of telemedicine entrepreneurship, identify key contributors, and highlight underexplored areas, thus contributing to a more holistic understanding of the field [15]. The research question is how to map and analyze the global trends of telemedicine entrepreneurship. Therefore, this study aims to map and analyze the global trends of telemedicine entrepreneurship. This research contribution identifies telemedicine trends, maps, gaps, and opportunities, offering practical insights to enhance healthcare accessibility, efficiency, and sustainability.

2 Materials and Method

This bibliometric analysis studied the intellectual structure in telemedicine and entrepreneurship literature using the Scopus database, known for its comprehensive coverage [2, 16]. The data search was conducted in July 2024, using phrases like "telemedicine" and "entrepreneurship" in the title, abstract, and keyword sections. The search query (TITLE-ABS-KEY (telemedicine* OR telehealth* OR telecare* OR "virtual care*" OR "remote healthcare*") AND TITLE-ABS-KEY (entrepreneur* OR startup OR SMEs)) yielded 165 articles selected for analysis from 1994 to 2023. Subsequently, the search results were then exported into a CSV dataset format [23]. Microsoft Excel visualizes the data through descriptive summary statistics and annual production performance.

The document selection process followed a bibliometric research workflow to obtain 165 relevant documents focusing on the intersection of telemedicine and entrepreneurship in business and management. This study used a period from 1994 to 2023 and included telemedicine and entrepreneurship as inclusion criteria for the data analyzed. Fig. 1 shows the workflow for selecting documents and bibliometric analysis.

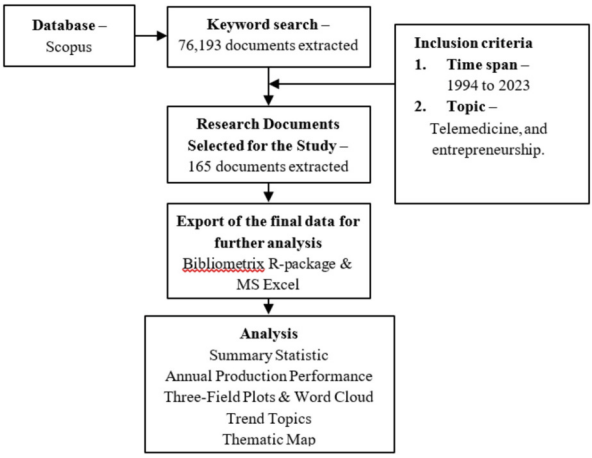


Fig. 1. Bibliometric analysis workflow

Data analysis was conducted using the Biblioshiny R package and Microsoft Excel. The R Biblioshiny package was used to visualize the data through descriptive summary statistics, three-field plots, word clouds, trend topics, and thematic maps [23]. Microsoft Excel visualizes the data through descriptive summary statistics and annual production performance. Research mapping analysis can be seen in summary statistics, three-field plots, word clouds, and thematic maps in telemedicine entrepreneurship [24]. Research trend analysis also included annual production performance and trend topics to track publication progress and identify emerging trends and research areas in telemedicine entrepreneurship. The analysis process was a systematic computation procedure [23]. The data obtained was compiled and validated.

3 Result

3.1 Summary Statistics

The research identified 165 documents from 139 sources published between 1994 and 2023, as seen in Table 1. The annual growth of 12.18% indicates an increased interest in this topic. The mean age of the documents was 6.39 years, with 11.91 citations per document, highlighting its impact on entrepreneurship. The research involved 4938 references and 581 authors, with 23.64% being international collaborators, indicating global attention and the importance of cross-border cooperation. Scientific articles dominated the document types, strongly focusing on scientific research.

Table 1. Summary statistics

Description	Results	Description	Results
Timespan	1994:2023	AUTHORS	
Sources	139	Authors	581
Documents	165	Authors of single-authored documented	32
Annual growth rate	12.18		
Document average age	6.39	Authors collaboration	
Average citations per doc	11.91	Single-authored documents	34
References	4938	Co-authors per document	3.76
		International co-authorships%	23.64
Document types			
Article	85		
Book chapter	11	Document contents	
Conference paper	30	Keywords plus (ID)	15.18
Conference review	5	Author's keywords (DE)	551
Letter	3		
Note	4		
Review	25		
Short survey	2		
Review	4		

3.2 Annual Production Performance

An analysis of annual production performance on telemedicine entrepreneurship from 1994 to 2023 shows a significant increase in research related to this topic, as seen in Fig. 2. In the beginning, research publications were few, signaling the early exploratory stage of telemedicine entrepreneurship. However, since 2010, the quantity of publications has risen markedly, peaking in 2023 with 28 publications, reflecting growing interest.

Fig. 2. Annual scientific production of telemedicine in entrepreneurship

3.3 Three-Field Plot and Word Cloud

Based on the three-field plot and word cloud analysis of telemedicine entrepreneurship (1994–2023) research, there is a significant correlation between citations, authors, and keywords in Scopus documents. The three-field plot illustrates the relationship between citations, authors, and keywords, highlighting the research's core themes and influential figures, as seen in Fig. 3. Authors such as Dal Mas F, Biancuzzi H, and Bednarova P are frequently cited, highlighting their essential contributions to the field. Meanwhile, the word cloud, as seen in Fig. 4, emphasizes the growing research interest and activity with keywords such as "telemedicine," "entrepreneurship," "digital health," "e-health," "telehealth," and "pandemic," indicating a growing focus on telemedicine technologies and supporting policies in healthcare and business strategies.

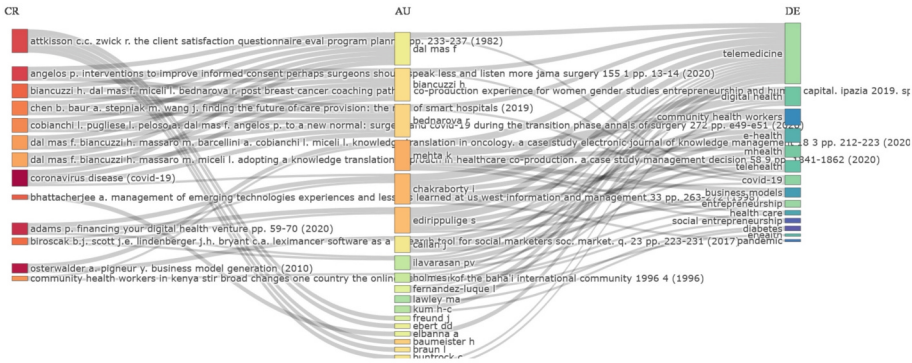


Fig. 3. Three-field plot of telemedicine in entrepreneurship

Fig. 4. Word cloud of telemedicine in entrepreneurship

3.4 Trend Topic

Based on a trend topics analysis of telemedicine entrepreneurship from 1994 to 2023, this study shows a significant evolution in research themes with a shifting focus over time. Terms such as "pandemic," "telehealth," "entrepreneurship," "public health," and "telemedicine" are becoming increasingly prominent, as seen in Fig. 5. The increased use of terms such as "pandemic" and "coronavirus disease 2019" post-2020 reflects the research response to the COVID-19 pandemic and the emerging role of telemedicine in sustaining healthcare. The rise of the terms "telemedicine" and "telehealth" signify the significant role of technology in expanding accessibility to and improving healthcare quality.

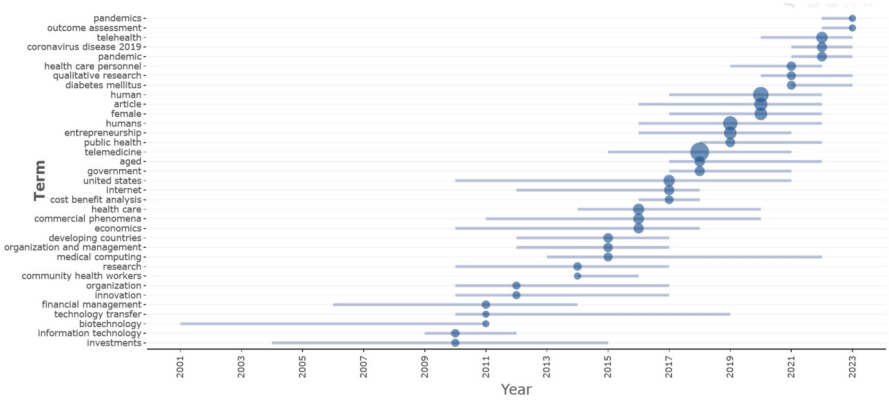


Fig. 5. Trend topics of telemedicine in entrepreneurship

3.5 Thematic Map

Based on the thematic map, the image identifies the main themes in telemedicine entrepreneurship research, with the position of the themes in four quadrants: motor themes, niche themes, essential themes, and emerging or declining themes, as seen in Fig. 6. Topics such as "entrepreneurship," "telehealth," and "health care" between the motor themes and basic theme positions indicate high relevance. They are frequently cited, reflecting the stability and dominance of these topics in entrepreneurship research. Topics of "telemedicine" and "human interaction" are between motor and niche themes, depicting growing research but more focused on specific or detailed aspects. Research topics such as "satellite communication systems" and "agriculture" are emerging or declining themes coverage, indicating a trend of innovation growth or declining interest and relevance. Essential themes support the structure and sustainable development of telemedicine entrepreneurship, with the overall theme forming the direction and basis of research based on technological innovation and health service improvement.

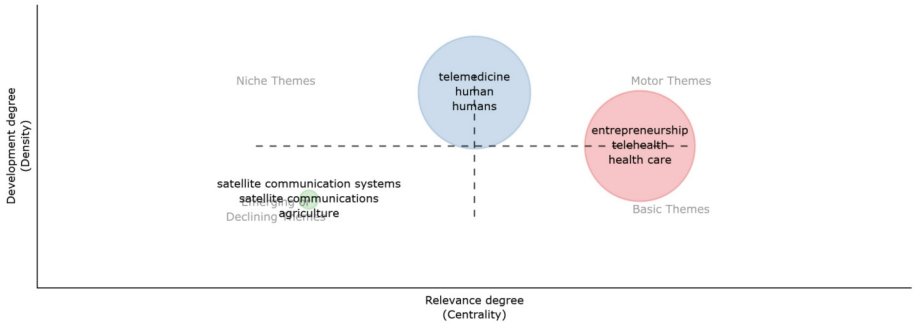


Fig. 6. Thematic maps of telemedicine in entrepreneurship

4 Discussion

This study uses bibliometric methods to explore trends and maps of telemedicine entrepreneurship research. One hundred sixty-five publications were analyzed as works of researchers from various countries indexed by Scopus published from 1994 to 2023. This summary statistics analysis shows a growing interest in telemedicine as part of entrepreneurship, with technological and policy support enabling innovation in healthcare [25]. This paradigm shift makes healthcare businesses increasingly dependent on telemedicine to optimize business operations and strategies [26]. Adapting to technological advancements becomes essential in planning effective business strategies, which align with the entrepreneurial trend of focusing on telemedicine solutions to improve innovation and health outcomes [27].

The annual production performance analysis identified that the increase in telemedicine entrepreneurship research is driven by advances in telemedicine technology, which open opportunities for innovation in business and healthcare strategies [28].

Increased accessibility and responses to international health issues, such as COVID-19, are accelerating the adoption of telemedicine for healthcare efficiency. The fluctuating dynamics of research indicate adaptation to new technologies and changing business strategies in the digital age [29].

The bibliometric analysis of telemedicine entrepreneurship reveals a significant increase in research activity and collaboration among scholars, as evidenced by the prominent authors and keywords identified in the three-field plot and word cloud. Notably, the contributions of authors such as Dal Mas F, Biancuzzi H, and Bednarova P underscore the growing academic interest in this field [29], paralleling findings from Dhyani et al., which highlight the obstacles and enablers for telemedicine adoption in low-income settings, emphasizing the role of local healthcare entrepreneurship [30]. Similarly, Kumar et al. discuss how social enterprises can drive innovation in healthcare, aligning with the emerging themes of digital health and entrepreneurship identified in our analysis [31]. Furthermore, Mishra and Pandey provide a comprehensive overview of global entrepreneurship trends in healthcare, reinforcing that telemedicine is critical for future research and entrepreneurial activity [32]. This synthesis of findings illustrates the dynamic interplay between telemedicine, entrepreneurship, and global health challenges, particularly in the context of recent technological advancements and the COVID-19 pandemic.

The trend topic analysis of telemedicine entrepreneurship from 1994 to 2023 indicates a notable evolution in research themes, particularly regarding the COVID-19 pandemic. The prominence of terms such as "pandemic," "telehealth," and "entrepreneurship" reflects an urgent shift in focus towards leveraging technology to enhance healthcare access and quality. This aligns with findings by Khandelwal et al., who emphasize the entrepreneurial prospects in digital healthcare that emerged post-COVID-19, particularly in developing countries [33]. However, Nicolau et al. focus primarily on the perspectives of Romanian entrepreneurs regarding digital sustainable entrepreneurship, which may not directly address the broader implications of the pandemic on telemedicine [34]. Furthermore, Lite et al. explore the relationship between health information technology and venture capital funding, underscoring the financial backing necessary for advancing telemedicine initiatives [35]. These studies corroborate the growing appreciation of telemedicine as an essential element of modern healthcare, driven by technological advancements and the pressing demands of public health crises.

The thematic map analysis of telemedicine entrepreneurship research reveals a clear delineation of themes, with "entrepreneurship," "telehealth," and "health care" positioned as motor and essential themes, indicating their high relevance and frequent citation in the literature. This finding is consistent with Liu's analysis, which identifies similar core themes in telemedicine research, emphasizing the importance of technology in healthcare delivery [26]. In contrast, Waqas et al. highlight the increasing publication output in telemedicine but do not specifically address the thematic evolution, thereby missing the nuanced shifts in focus towards niche areas like "human interaction" and emerging themes such as "satellite communication systems" [36]. Furthermore, Xie et al. explain and document the exponential growth of telemedicine research in Canada, which aligns with global trends but lacks a comprehensive thematic mapping that captures the interplay of various topics [37]. Our analysis fills this gap by illustrating how essential

themes support the sustainable development of telemedicine entrepreneurship, thereby providing a structured framework for future research that addresses both technological innovation and health service improvement, which previous studies have not fully explored.

The primary limitation of this research lies in its exclusive reliance on data from the Scopus database, potentially overlooking relevant studies published in other academic databases, which may limit the comprehensiveness of the findings. Additionally, while bibliometric analysis provides a quantitative overview of research trends, it lacks the depth and contextual understanding that qualitative methods might offer, such as insights into barriers or motivations behind telemedicine entrepreneurship. Though effective, the study's use of R Biblioshiny and Excel for data visualization may restrict the analytical depth compared to more sophisticated data analytics tools.

5 Conclusion

This study has answered the objective of mapping and analyzing global trends in telemedicine entrepreneurship. This study reveals significant interest worldwide in leveraging telemedicine as an innovation strategy in entrepreneurship through a bibliometric analysis of 165 documents collected from Scopus from 1994 to 2023. This trend underscores the increasing interest and recognition of the significant impact of telemedicine technology on the development of entrepreneurship and healthcare strategies based on annual production performance. In addition, the trend topics refer to innovation, entrepreneurship, telehealth, and digital health. The three-field plot analysis highlights telemedicine applications with an international and collaborative focus on entrepreneurship by involving frequently cited authors, keywords, and references. The word cloud groups important research topics or terms such as telemedicine, entrepreneurship, digital health, and pandemic related to the main research themes. In addition, the thematic map categorizes research themes into motor, basic, specialized, and emerging areas to identify focus topics. Themes such as entrepreneurship, telehealth, and healthcare were the main driving motor themes. Future research on telemedicine entrepreneurship can explore, at the micro level, how individual characteristics, cognitive processes, and digital competencies like adaptability and resilience influence business success. At the meso level, it can examine ecosystem dynamics, including interactions among healthcare institutions, technology providers, and regulatory frameworks, to identify factors supporting innovation and sustainability. At the macro level, research can analyze the impact of policies, regulations, and economic incentives on telemedicine's growth and assess its socioeconomic effects on patient outcomes, care delivery, and public health. This study's implication is the importance of telemedicine in driving innovation and efficiency in healthcare entrepreneurship. These findings are important as a tool for entrepreneurs to utilize the potential of digital technology to create flexible and efficient business strategies, meet patient needs, and deliver innovation in healthcare services. Telemedicine will ensure a plan for more flexible digital solutions amidst changing market dynamics.

Acknowledgments. This manuscript is a publication output from an undergraduate thesis track on scientific articles in reputable international publications at the Entrepreneurship Department,

Bina Nusantara University, Malang Campus. The authors would like to thank the Entrepreneurship Department and Research and Technology Transfer Office, Bina Nusantara University, for supporting the research ecosystem, facilities, and publication funding.

Author Contributorship. **Noor Anisa Nabila:** Data Curation, Formal Analysis, Investigation, Methodology, Resources, Software, Visualization, Writing–Original Draft Preparation, Writing–Review and Editing; **Agung Purnomo:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Resources, Validation, Funding Acquisition, Project Administration, Supervision, Writing–Review and Editing.

Data Availability. **Harvard Dataverse:** "Dataset of Telemedicine Entrepreneurship (1994–2023)." https://doi.org/10.7910/DVN/WGTWLW (Nabila, N.A., Purnomo, A., 2024).

References

1. Zakiyyah MI, Pramono A, Santosa RG, Purnomo A, Maulana FI (2023) Agustina, I.A.: haters fat: Mobile health application to help the productive young generation achieve a healthy life. In: 2023 10th international conference on ICT for smart society (ICISS). IEEE, Bandung, pp. 1–6. https://doi.org/10.1109/ICISS59129.2023.10291913
2. Wilson LS, Maeder AJ (2015) Recent directions in telemedicine: review of trends in research and practice. Healthc Inform Res 21:213. https://doi.org/10.4258/hir.2015.21.4.213
3. Stewart MK, Allison MK, Grant Hunthrop MS, Marshall SA, Cornell CE (2023) Outcomes research on telemedicine-delivered gender-affirming health Care for Transgender Youth is Needed now: a call to action. Transgender Health 8:1–5. https://doi.org/10.1089/trgh.2021.0063
4. Garavand A, Aslani N, Behmanesh A, Shams R (2022) Telemedicine in lung cancer during COVID-19 outbreak. J Educ Health Promot 11:348. https://doi.org/10.4103/jehp.jehp_50_22
5. Mohammadzadeh N, Rezayi S, Saeedi S (2023) Telemedicine for patient Management in Remote Areas and Underserved Populations. Disaster Med Public Health Prep 17:e167. https://doi.org/10.1017/dmp.2022.76
6. Fassas S, Cummings E, Sykes KJ, Bur AM, Shnayder Y, Kakarala K (2021) Telemedicine for head and neck cancer surveillance in the COVID-19 era: promise and pitfalls. Head Neck 43:1872–1880. https://doi.org/10.1002/hed.26659
7. Aghdam MRF, Vodovnik A, Hameed RA (2019) Role of telemedicine in multidisciplinary team meetings. J Pathol Inform 10:35. https://doi.org/10.4103/jpi.jpi_20_19
8. Aliberti GM, Bhatia R, Desrochers LB, Gilliam EA, Schonberg MA (2022) Perspectives of primary care clinicians in Massachusetts on use of telemedicine with adults aged 65 and older during the COVID-19 pandemic. Prev Med Rep 26:101729. https://doi.org/10.1016/j.pmedr.2022.101729
9. Bowen ME, Bosworth HB, Roumie CL (2013) Blood pressure control in a hypertension telemedicine intervention: does distance to primary care matter? J Clin Hypertens 15:723–730. https://doi.org/10.1111/jch.12172
10. Gmunder KN, Ruiz JW, Franceschi D, Suarez MM (2024) Demographics associated with US healthcare disparities are exacerbated by the telemedicine surge during the COVID-19 pandemic. J Telemed Telecare 30:64–71. https://doi.org/10.1177/1357633X211025939
11. Dal Mas F, Biancuzzi H, Bednarova R, Miceli L (2022) A gender perspective on telemedicine. Early results from the National Cancer Institute of Aviano experience. In: SIDREA series in accounting and business administration. Springer, Cham, pp 73–83. https://doi.org/10.1007/978-3-030-89412-2_6

12. Sechrist E et al (2020) Patients' perception of telemedicine in a large urban Inner-City emergency department: a cross-sectional survey. Cureus 12. https://doi.org/10.7759/cureus.11091
13. Morgan A, Goodman D, Vinagolu-Baur J, Cass I (2022) Prenatal telemedicine during COVID-19: patterns of use and barriers to access. JAMIA Open 5:1–6. https://doi.org/10.1093/jamiaopen/ooab116
14. Kaur K, Kapur B (2024) Trends in remote healthcare and the future of telemedicine: a bibliometric analysis. In: Advances in economics, business and management research. Atlantis Press, pp 377–394. https://doi.org/10.2991/978-94-6463-544-7_25
15. Aria M, Cuccurullo C (2017) Bibliometrix : an R-tool for comprehensive science mapping analysis. J Informetr 11:959–975. https://doi.org/10.1016/j.joi.2017.08.007
16. Rousseau R, Egghe L, Guns R (2018) Becoming metric-wise: a bibliometric guide for researchers. Elsevier Ltd, Cambridge
17. Sahoo S, Sahoo J, Kumar S, Lim WM, Ameen N (2023) Distance is no longer a barrier to healthcare services: current state and future trends of telehealth research. Internet Res 33:890–944. https://doi.org/10.1108/INTR-10-2021-0774
18. Kozikowski A et al (2019) Care team perspectives and acceptance of telehealth in scaling a home-based primary care program: qualitative study. JMIR Aging 2:e12415. https://doi.org/10.2196/12415
19. Theodore BR et al (2015) Transaction cost analysis of in-clinic versus telehealth consultations for chronic pain: preliminary evidence for rapid and affordable access to interdisciplinary collaborative consultation. Pain Med 16:1045–1056. https://doi.org/10.1111/pme.12688
20. Saberi P, Yuan P, John M, Sheon N, Johnson MO (2013) A pilot study to engage and counsel HIV-positive African American youth via telehealth technology. AIDS Patient Care STDs 27:529–532. https://doi.org/10.1089/apc.2013.0185
21. Ziegler E et al (2024) Response and innovations of advanced practice nurses during the COVID-19 pandemic: a scoping review. Int Nurs Rev 71:250–275. https://doi.org/10.1111/inr.12884
22. Jack SM et al (2021) Recognising and responding to intimate partner violence using telehealth: practical guidance for nurses and midwives. J Clin Nurs 30:588–602. https://doi.org/10.1111/jocn.15554
23. Purnomo A, Septianto A, Sutiksno DU, Hikmawan MI, Kumalasari RD (2020) Technopreneur publication: a bibliometric analysis (2000–2019). In: 2020 International conference on information management and technology (ICIMTech). IEEE, pp 521–526. https://doi.org/10.1109/ICIMTech50083.2020.9211111
24. Purnomo A, Firdaus M, Maulana FI, Triangga B, Hikmawan MI, Hikmah ZTN (2024) Business intelligence in healthcare: a review of knowledge structures and level of analysis. In: Lecture notes in electrical engineering. Springer, Singapore, pp 505–517. https://doi.org/10.1007/978-981-97-1463-6_35
25. Valdes D, Alqazlan L, Procter R, Dale J (2022) Global evidence on the rapid adoption of telemedicine in primary care during the first 2 years of the COVID-19 pandemic: a scoping review protocol. Syst Rev 11:124. https://doi.org/10.1186/s13643-022-01934-3
26. Liu P, Wang F, Xu W, Li Y, Li B (2023) Trends and frontiers of research on telemedicine from 1971 to 2022: a scientometric and visualisation analysis. J Telemed Telecare 29:731–746. https://doi.org/10.1177/1357633X231183732
27. Del Carpio-Delgado F et al (2023) Telemedicine in Latin America: a bibliometric analysis. EAI Endorsed Trans Pervasive Health Technol 9. https://doi.org/10.4108/eetpht.9.4273
28. King CA, Cook R, Korthuis PT, McCarty D, Morris CD, Englander H (2022) Expanding inpatient addiction consult services through accountable care organizations for Medicaid Enrollees: a Modeling study. J Addict Med 16:570–576. https://doi.org/10.1097/ADM.0000000000000972

29. Dal Mas F, Biancuzzi H, Marseglia GR, Bednarova R, Cobianchi L, Miceli L (2022) Doctor@home: new perspectives on telemedicine for women during the COVID pandemic. In: Digital innovation for healthcare in COVID-19 pandemic: strategies and solutions. Elsevier, pp 53–66. https://doi.org/10.1016/B978-0-12-821318-6.00007-4
30. Dhyani VS et al (2023) Barriers and facilitators for the adoption of telemedicine services in low-income and middle-income countries: a rapid overview of reviews. BMJ Innov 9:215–225. https://doi.org/10.1136/bmjinnov-2022-001062
31. Kumar P, Paton C, Kirigia D (2016) I've got 99 problems but a phone Ain't one: electronic and mobile health in low and middle income countries. Arch Dis Child 101:974–979. https://doi.org/10.1136/archdischild-2015-308556
32. Mishra A, Pandey N (2023) Global entrepreneurship in healthcare: a systematic literature review and bibliometric analysis. Glob Bus Organ Excell 42:9–21. https://doi.org/10.1002/joe.22193
33. Khandelwal R, Kolte A, Rossi M (2022) A study on entrepreneurial opportunities in digital health-care post-Covid-19 from the perspective of developing countries. Foresight 24:527–544. https://doi.org/10.1108/FS-02-2021-0043
34. Nicolau C, Nichifor E, Munteanu D, Bărbulescu O (2022) Decoding business potential for digital sustainable entrepreneurship: what Romanian entrepreneurs think and do for the future. Sustainability 14:13636. https://doi.org/10.3390/su142013636
35. Lite S, Gordon WJ, Stern AD (2020) Association of the Meaningful use Electronic Health Record Incentive Program with Health Information Technology Venture Capital Funding. JAMA Netw Open 3:e201402. https://doi.org/10.1001/jamanetworkopen.2020.1402
36. Waqas A, Teoh SH, Lapão LV, Messina LA, Correia JC (2020) Harnessing telemedicine for the provision of health care: bibliometric and Scientometric analysis. J Med Internet Res 22:e18835. https://doi.org/10.2196/18835
37. Xie JS et al (2022) Publication trends in telemedicine research originating from Canada. Healthc Manag Forum 35:153–160. https://doi.org/10.1177/08404704211070240

Biomechanical Simulation and Experimental Study for Total Hip Arthroplasty Placement Tools: Femoral Broach and Acetabular Reamer

Dita A. Mayasari[1,2(✉)], Nonik S. Ayati[1], Nisa R. Amitasari[1], Talitha Asmaria[3], Menik Dwi Kurniatie[1,2], Joko Triwardono[3], and Muhammad S. Utomo[3,4]

[1] Department Biomedical Engineering, Dian Nuswantoro University, Semarang, Indonesia
mayasari.dita@dsn.dinus.ac.id
[2] Center for Medical Technology Innovation (CEMTI), Dian Nuswantoro University, Semarang, Indonesia
[3] National Research and Innovation Agency (BRIN), Jakarta, Indonesia
[4] Department of Biomedical Engineering, University of Melbourne, Victoria, Australia

Abstract. Total hip arthroplasty (THA) is an orthopedic surgery to solve hip disorders. During placement to the patient, THA, which consists of the hip stem and acetabular components, must follow the morphological profile of the patient. Plenty of studies have described the importance of patient-specific implant (PSI) for THA, both of how to design and implement it, but no one published how to adjust the attached installation tools. The THA installation tools are a femoral broach for installing the hip stem and an acetabular reamer for installing acetabular components. To the best of our knowledge, this is the first study that develops the THA installation tools, consisting of a femoral broach and acetabular reamer, that follow the hip and femur bone's patient profiles. First, the designed patient-specific installation tools of the femoral broach and acetabular reamer are computed using finite element analysis (FEA) to understand the biomechanical responses when it is applied. Second, through self-design experimental testing using a 3D print prototype, compatibility between all components is analyzed. From FEA, the maximum Von Mises Stress (VMS) in the femoral broach is 76.206 MPa and the acetabular reamer reached 150.7 MPa, which both are lower than the SS316L yield strength value 170–300 MPa and indicates the designed femoral broach and acetabular reamer are safe to be manufactured. Our experimental tests show compatibility between all components while broaching and reaming. The successful design and experimental studies of THA installation tools could significantly help industries, particularly in low-income countries, to independently manufacture medical devices.

Keywords: Hip joint replacement · Hip implant tools · Stainless steel · Von Mises stress

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 53–70, 2026.
https://doi.org/10.1007/978-981-95-9233-3_5

1 Introduction

Total hip arthroplasty (THA) is one of the most common surgeries worldwide as well as one of the significant advancements in the surgery of orthopedic disorders [1]. However, in performing THA, post-surgical complications remain a problem for many patients [2]. Advances in implant design, the use of advanced protocols to control infection, and the use of minimally invasive procedures have improved THA outcomes [3]. However, possible causes of THA failure include dislocation and implant misalignment so these need to be addressed to improve THA results [4]. Femoral stem misalignment in THA can have a detrimental impact on the outcome. Poor intraoperative preparation of the femur is an important cause of malalignment. The position of the femoral stem is influenced by the design of the instruments that used to prepare the femur [5]. THA procedure involves several crucial surgical instruments, including the femoral broach and acetabular reamer [6]. During replacement surgery, the equipment is used to prepare the bone surface and facilitate the insertion of implant components [7]. A femoral broach is used to prepare the femur for the placement of the femoral component. The femoral broaches are available in various sizes and designs, and instruments are usually pointed in shape with sharp edges that help shape the femoral canal [8]. While the acetabular reamer is used to prepare the acetabulum for the placement of the acetabular component, the design comes in various sizes and shapes. They are typically attached to a surgical drill for a reaming system and are available in manual and advanced versions [9].

Research regarding surgical instruments in THA, particularly the femoral broach and the acetabular reamer is very limited. Preutenborbeck et al. conducted research to determine the efficiency of broaching during the preparation of the femoral cavity using a femoral broach [10]. In that study, cadaveric cortical bone was used, cancellous bone used foam, and femoral broach used metal material. The aim was to characterize impactions during broaching [10]. In their studies, there is no information on how to develop designs for THA surgical tools. Surgical instruments are typically made from certain metals such as stainless steel, cobalt-chrome, or titanium [11]. While other materials like cobalt-chrome and titanium are also used, stainless steel remains a prevalent choice in the manufacturing of these orthopedic instruments [12]. An example product from Enztec produces acetabular reamers using stainless steel. Meanwhile, an example of an industry that manufactures femoral broach using stainless steel material is AZ Ortho. In the configuration of acetabular reamers, the common size can range from 40 mm to 80 mm or more, and the femoral broach may range from 9 mm to 17 mm or more [9]. A common material used in the production of biomedical implants and surgical instruments is stainless steel 316L (SS316L), SS316L is a strong steel that has excellent corrosion resistance, high-temperature strength, and good mechanical properties, “L” stands for low carbon which indicates this grade has a lower carbon content than standard 316 stainless steel. The low carbon content helps to improve corrosion resistance [13].

Human anatomy, including the acetabulum and femur, varies greatly between individuals, and there is no standard size that applies to people from all over the world [14]. The size and shape of a person's acetabulum and femur can vary greatly based on factors such as gender, genetics, height, leg length, sex, and individual anatomical variations [15]. Inaccurate reamer and broach size designs can risk increased wear of the joint surfaces and breaking and deforming easily [16, 17]. In the THA procedure, the efficiency, accuracy, and biomechanical compatibility of these tools are critical to ensure the success and long-term stability of the hip implant. As the field of orthopedic surgery continues to evolve, the quest for innovation in THA procedures is becoming increasingly important. Such procedures depend not only on the skill of the surgeon but also on the effectiveness of the tools used, especially the femoral broach and acetabular reamer [3]. Overall previous papers have declared the evaluation of the femoral broach application, the rough size of both femoral broach and acetabular reamer, and the used material for manufacturing. However, detailed information on how to design from the patient-CT data, in the case of personalized need, is not available.

The aim of this study is to improve the design of the acetabular reamer and femoral broach through finite element analysis (FEA) computational simulation testing to understand and determine the performance of the instruments under loading conditions, select appropriate materials, and evaluate reaming and broaching results that involves stress assessment before printing [18]. To improve the performance of the design, experimental testing was conducted to determine the performance of the implant placement tool to mimic the actual conditions. This paper is expected to provide implications on the manufacturing processes that need to be adjusted to accommodate the production of personalized acetabular reamers and femoral broaches, ensuring proper instrument production based on FEA testing and experimental testing. Being able to self-manufacture the implant and follow surgical tools is significantly helping a country, particularly low-middle income countries, such as Indonesia, to be more independent providing access to more affordable medical devices.

2 Materials and Methods

The data used to design the femoral broach and acetabular reamer were obtained from measurements of the femur bone using parameters including femoral head offset (FHO), femoral head diameter (FHD), femoral neck length (FNL), mediolateral width (MLW), and neck shaft angle (NSA) are shown in Fig. 1. Data was obtained from a dual source CT-Scan 128 slices belonging to Cipto Mangunkusumo Hospital Jakarta with permission from a radiologist and orthopedic expert. The data is in the form of a CT of the pelvis or lower abdomen. The patient is female, weighs 61 kg, has a height of 165 cm, and is 49 years old [19]. The data used to design the femoral broach and acetabular reamer refers to parameters including femoral cap offset (FHO), femoral cap diameter (FHD), femoral neck length (FNL), mediolateral width (MLW), and neck. Axis angle (NSA) is shown in Fig. 1.

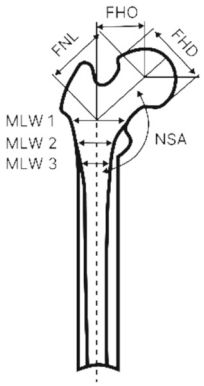


Fig. 1. Geometric parameters of the femur bone.

This section outlines the methodology used to investigate THA placement devices, specifically focusing on the design and analysis of the acetabular reamer and femoral broach. This study employed the ANSYS software as a finite element (FE) solver to simulate the 3D FE model. The basic workflow used in the FE model simulation uses explicit dynamics to show broaching simulations using a femoral broach on cancellous bone and reaming simulations using an acetabular reamer on the acetabular bone. The implementation procedure is shown in the Fig. 2.

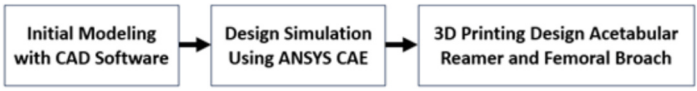


Fig. 2. The implementation procedure.

2.1 Femoral Broach

Modeling of femoral broach surgical instruments with adjustments to the morphology of the femur and femoral stem was designed in the Autodesk Fusion 360 software, which is then saved in the Standard for the Exchange of Product (STEP) file format. There are several parameters on the femoral broach, including mediolateral width 1 (MLW1), mediolateral width 2 (MLW2), mediolateral width 3 (MLW3), length of the femoral broach, length of the handle, diameter of the handle, the distance between broaches, and broach depth shown in Fig. 3. Figure 4 shows sections of the cortical bone and cancellous bone in the femur as well as the femoral broach component. Femoral broach will later broach the cancellous bone by leaving a distance of 3–4 mm from the cortical bone. The femoral broach has a length of 92.3 mm, an MLW1 of 21.32 mm, an MLW2 of 15.85 mm, an MLW3 of 8.76 mm, an FHO of 30.72 mm, an FNL of 22.361 mm, an handle length of 15 mm, a handle diameter of 6 mm, a distance between each broach and the broach depth of 2 mm.

The computer-aided design (CAD) file of the femur bone and femoral broach were imported into the software of ANSYS workbench for preparing the mesh, defining the contact, and setting the boundary conditions. There are five stages in using the finite element analysis (FEA) using ANSYS software with explicit system dynamics analysis including engineering data, geometry, model, setup, and solution. This engineering data will include the material properties of cortical bone, cancellous bone, and femoral broach as shown in Table 1. The femoral broach is arranged to be an SS316L material.

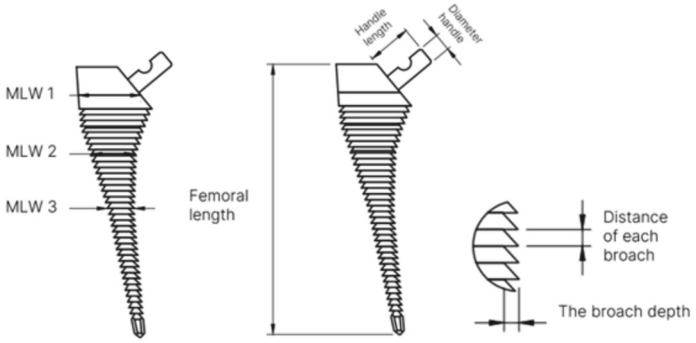


Fig. 3. Femoral broach size parameters.

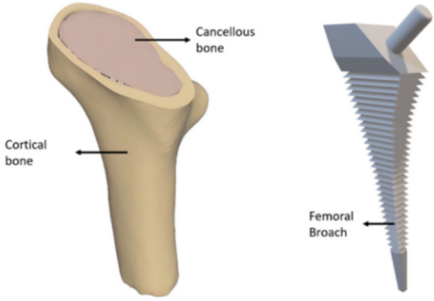


Fig. 4. The cancellous and cortical bone in the femur and the femoral broach.

Table 1. Property material and mechanical property [9]

Properties	Material properties		
	Cortical bone	Cancellous bone	SS316L
Density (kg/m^3)	1700–2000	580–1700	8000
Young's modulus (MPa)	3800–11700	10–500	193
Poisson ratio	0.3	0.3	0.3
Compressive strength (MPa)	88–164	23	235
Tensile strength (MPa)	82–114	10–20	170
Yield strength (MPa)	–	–	170–300

In the geometry stage of the femoral broach simulation, the geometry import will be carried out in STEP format, and the position of the cancellous bone, cortical bone, and femoral broach will be determined using the software ANSYS Design Modeler. The use of the STEP format is because the file size is smaller and can store more complete information about the model, including topology, parameters, and structure. This makes it possible to avoid errors in data processing. The stage of 3D modeling consists of connecting the geometry with the material properties used, connections, and mesh. In connections, it will select body contacts and target contacts on the object. The femoral broach simulation uses a body contact setting in the form of a femoral broach and a target contact on the cancellous bone with a frictionless type. Meanwhile, the mesh uses a tetrahedral element shape with four nodes. The choice of "tet4" is due to the complex geometry of the femoral broach and femur and due to the maximum limits of ANSYS student license nodes and elements. The setup for the femoral broach simulation will include the velocity, force, displacement, and fixed support values used. In the initial condition, the velocity was set at 3100 mm/s which was adjusted to the velocity from previous research that carried out experimental testing. Meanwhile, the displacement is − 200 mm and the force used is 260 N on the x-axis, they are calculated from the upper end of the femoral broach and it is shown in Fig. 5.

The preparation process for experimental testing includes 3D printing the cortical bone design, femoral broach design, femoral broach handle, and frame using PLA material and a 3D machine in the form of Kingroon KP3S. Meanwhile, for the cancellous bone, solid soap material is used which is melted and re-solidified into the 3D-printed cortical bone. After obtaining the 3D printing results and printing the cancellous bone with soap, a broaching process will be carried out using a femoral broach on the cancellous bone. The experimental test results will pay attention to whether the femoral broach can broach the cancellous bone and fit properly in the femur bone as shown in Fig. 6.

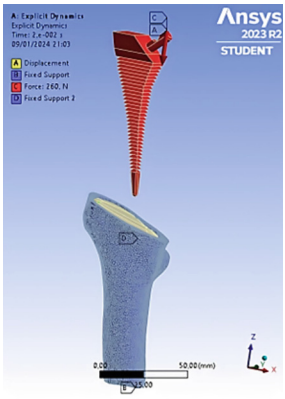


Fig. 5. Femoral broach setting for the FEA simulation.

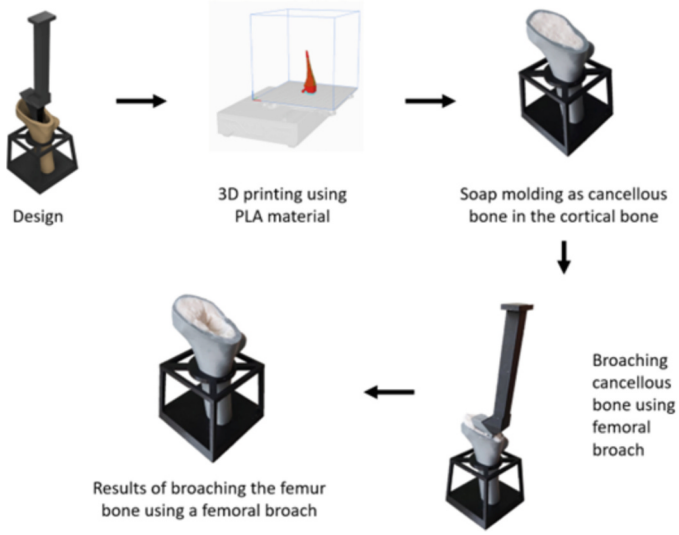


Fig. 6. The process of experimental testing on the femoral broach and the femur bone.

2.2 Acetabular Reamer

The acetabular reamer design is hemispherical in shape with an angle of 167.65°, a diameter of 42 mm, a height of 21 mm, with a total of 21 serrations of 4 mm diameter and 1mm thickness. A 46 mm long connecting rod was designed to connect the acetabular reamer to the drill. The design of the acetabular reamer is shown in the Fig. 6 is made with reference to the implant size of the acetabular component [19]. The acetabulum bone for FEA simulation uses an acetabulum bone with a size of 42 mm, the FEA simulation used a student license so there are limitations to the structural physics geometry, so a reconstruction of the acetabulum geometry is required, reconstruction of the acetabulum geometry is shown in Fig. 7 [20]. Design of the geometrical data that has been obtained

is then carried out using finite element analysis by computer-aided engineering (CAE) software of ANSYS workbench student license [20]. The first step is to import the geometry of the STEP file. The STEP format is a standard for representing 3D models and is widely used in computer-aided design (CAD) and product data exchange. It is an ISO standard (ISO 10303). The STEP files have the file extension ".stp" or ".step." [21].

Material assignment of acetabular reamer material can affect performance and durability so that it does not cause adverse reactions when in contact with the human body. This engineering data will include the material properties of acetabulum bone and acetabular reamer as shown in Table 2. The acetabular reamer is assumed to be SS316L material.

Meanwhile, in the acetabular reamer simulation as shown in Fig. 8, adjusting the movement the rotation speed was selected as 180 rpm with a Z-axis displacement into the acetabulum of 8 mm, the selection of the rotation speed was based on the drill machine specifications with the aim that the existing acetabulum could be eroded by the acetabular reamer, the displacement to the Z-axis given aims to provide contact of the acetabular reamer to the acetabulum. The application depends on several factors, one of which is the specifications of the drill machine tool, the load is applied as much as 2.5 N according to the specifications of the drill tool used in the experimental test. At the solution stage, the results of the analysis were known to be von Mises stress (VMS) on the femoral broach and the acetabular reamer. These results will show the distribution of stress contours on the femoral broach and acetabular reamer which is simulated using software in the form of an image where the value is represented by a gradient of blue for safe areas to red for the most critical conditions.

After the design passes the simulation test, it is printed using 3D printing. The 3D printed design is then inserted into the 3D pelvic bone acetabulum, where it determines that the specified acetabular reamer is appropriate. The acetabular component with the acetabular reamer and the acetabular reamer with the 3D pelvic bone having the correct size is the success parameter used to determine the results. The accuracy measured is not too small and not too big. After the experimental test of size suitability, the next experimental test of reaming was carried out, the 3D acetabular reamer was attached to a portable drilling machine and then drilled into an object in the form of a gypsum. Determination of success in the form of reaming results has a suitable hole cut when the acetabular component is attached to the hole.

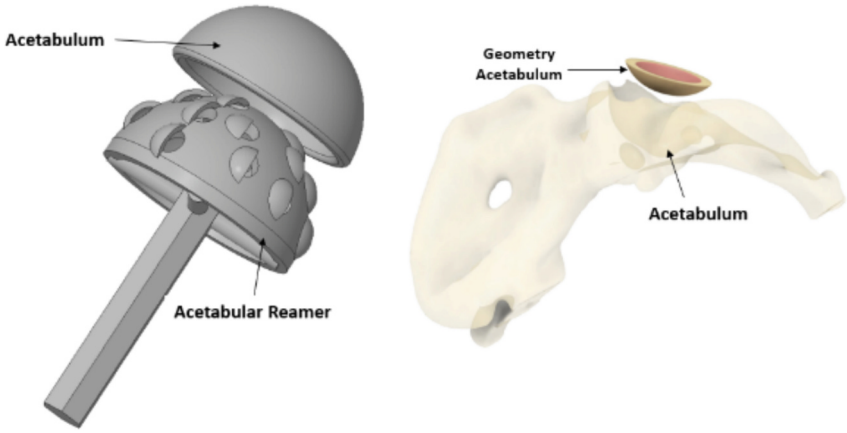


Fig. 7. Acetabular reamer components and acetabulum.

Table 2. Material and mechanical properties of acetabular reamer [10, 12].

Properties	Material properties	
	Acetabulum	SS316L
Density (kg/m^3)	348.3	8000
Young's modulus (MPa)	116.4	193
Poisson ratio	0.3	0.3
Compressive strength (MPa)	–	235
Tensile strength (MPa)	–	170
Yield strength (MPa)	–	170–300

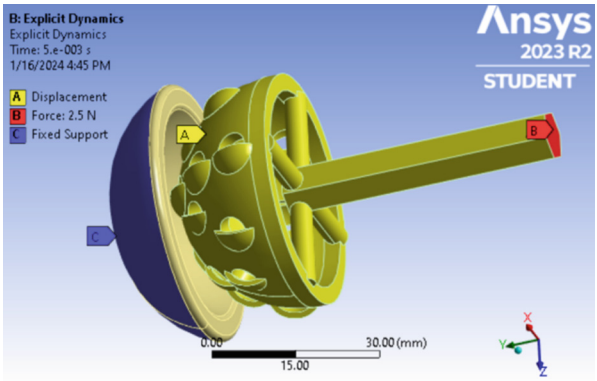


Fig. 8. Acetabular reamer motion limit setting in the FEA simulation.

2.3 Analysis Method

The analysis method used is the finite element method to evaluate the strength and stability of materials or structures under certain loads. In structural analysis, each element receives a force or load and produces a response through stress and strain. Each element will experience three-dimensional stress, including normal stress and shear in various directions. These stresses are usually calculated in terms of principal stresses (σ_1, σ_2, and σ_3) at each point in the element. Structural failure is determined from the equivalent stress using von Mises Theory. Von Mises stress (σ_{vm}) is calculated based on the principal stress of the element. The von Mises formula for equivalent stress is Eq. (1):

$$\sigma_{vm} = \sqrt{\frac{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2}{2}} \tag{1}$$

This equivalent stress is compared directly with the material's yield strength value. Structural failure is considered to occur if the von Mises stress exceeds the yield limit value [22]. A regression analysis was performed on the von Mises stress results to assess the potential for materials to experience plastic deformation or fail (failure) based on yield criteria and data convergence. Significant correlations between these parameters and stress levels were identified, suggesting critical factors that may contribute to material failure or deformation under load conditions. A multiple linear regression model might relate equivalent stress to design and load parameters by Eq. (2):

$$\sigma_{vonMises} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \epsilon \tag{2}$$

here, $\sigma_{vonMises}$ represents the von Mises stress (dependent variable), $x_1, x_2, \ldots, x_n$ are independent variables (e.g., material thickness, load magnitude), β values are coefficients indicating the strength of each relationship, and ϵ is the error term [23].

3 Results

3.1 Femoral Broach

The results of von Mises stress (VMS) on the femoral broach get a maximum value is 76.206 MPa and a minimum value is 0.39766 MPa as shown in Fig. 9. Meanwhile, the VMS results on cancellous bone get a maximum value of 13.584 MPa and a minimum value of 0.063302 MPa as shown in Fig. 10.

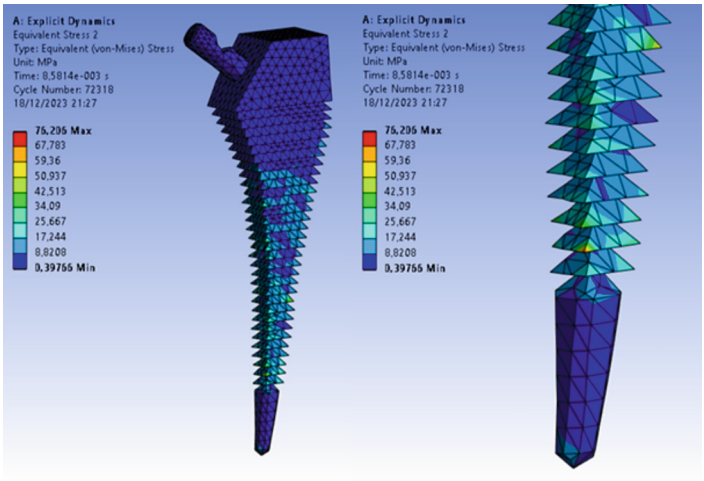


Fig. 9. The result VMS of a femoral broach.

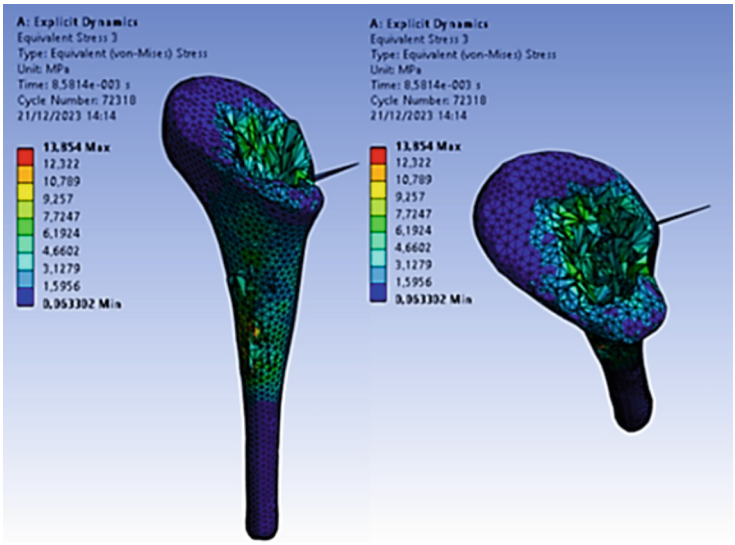


Fig. 10. The result VMS of cancellous bone.

The results of analysis using FEA to study the mechanical response of the femoral broach when applying load, force, and velocity to the cancellous bone show that the femoral broach design can be used to broach the cancellous bone. Based on these results, it is estimated that the part that will experience a fracture of the femoral broach is the lower cleat which has been marked in Fig. 9. The results of FEA simulation broaching on cancellous bone were 3D printed using resin to determine the results of broaching

carried out in FEA testing using ANSYS software. The results of the cancellous bone that has been broached are shown in Fig. 11.

Fig. 11. The result 3D prints cancellous bone after broaching.

Based on the Equivalent Stress graphic image from the Explicit Dynamics simulation in ANSYS, it was found that the simulation was carried out in a very short period, with a final time of around 8.5814×10^{-3} s. This often occurs in explicit analysis for fast events or dynamic loads. Figure 12 shows that the equivalent stress value experiences a gradual increase during the simulation. The stress tends to increase gradually and reaches a value of around 70 MPa at the end of the simulation. The graph shows some minor fluctuations around the peak value, but the trend is generally stable with a gradual increase. In this graph, although there are slight fluctuations, the equivalent stress trend is relatively stable and increases predictably. Small fluctuations in the graph could be due to inertial effects or local instability in the simulation, but these do not appear to be significant. This graph shows stable results and can be considered "convergent" in explicit dynamics, even though explicit analysis does not use iterative convergence criteria. The gradual increase of the equivalent stress without large spikes indicates that the material responds reasonably to the applied load in this simulation [24].

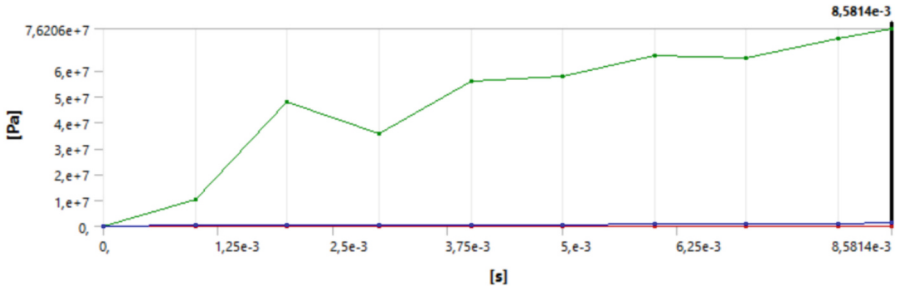


Fig. 12. Equivalent stress graphic image

The experimental testing process begins by 3D printing the design of the testing equipment in Fig. 13a and preparing for cancellous bone splinting using soap that is melted and then compacted again in the trabecular bone. The broaching process is shown in Fig. 13b and a comparison of the femur bone before and after the broaching process is

obtained in Fig. 14. The results of femoral broaching in the cancellous bone leave more than the same as 2 mm of cortical bone.

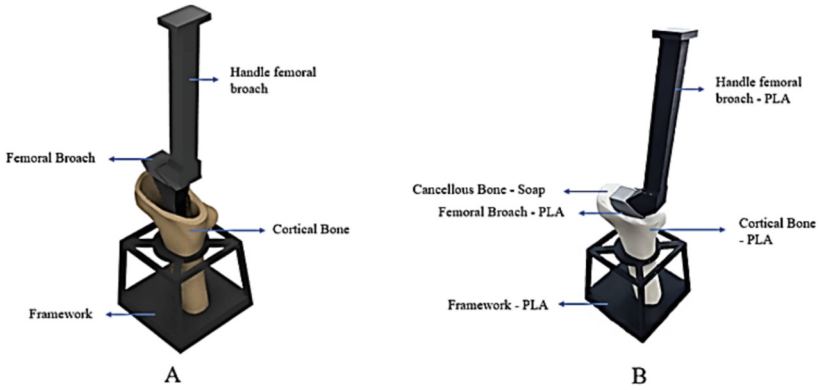


Fig. 13. Testing equipment (**a**) design of experimental testing equipment and (**b**) experimental testing.

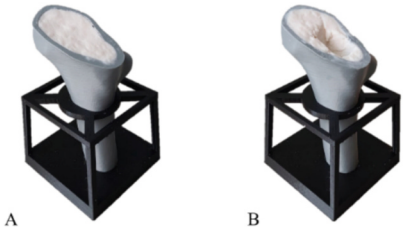


Fig. 14. Femur bone: (**a**) before broaching and (**b**) after broaching.

The choice of soap material takes into account the ratio of Young's modulus in soap and PLA materials, which is the same as the ratio of cancellous bone material and SS316L, that is 1:20. The comparison of Young's modulus soap of 15 MPa compared to young's modulus of PLA of 300 MPa is the same as Young's modulus of cancellous bone of 10 MPa compared to Young's modulus of SS316L of 200 MPa [25].

3.2 Acetabular Reamer

The result of the FEA simulation given in Fig. 15 shows a maximum VMS value of 150.7 MPa. Based on the FEA results, the design is not expected to experience permanent deformation or failure under 2.5 N loading conditions because the maximum VMS is lower than the yield strength of SS316L which is around 170–300 MPa. FEA results are displayed in color, with red representing the highest VMS values, orange and yellow indicating areas with medium VMS values, and blue and green indicating areas with low

VMS values. Figure 15 shows that the highest VMS loads occur because the acetabular reamer teeth receive most of the movement directly into the acetabulum.

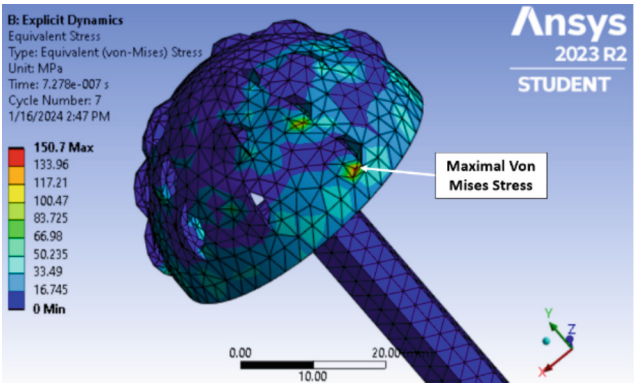


Fig. 15. Equivalent VMS results on the acetabular reamer.

Figure 16 describes the design results of the 3D printed acetabular reamer placed into the acetabulum showing appropriate dimensions that are not too large and not too small, as well as dimensional conformity to the acetabular shell.

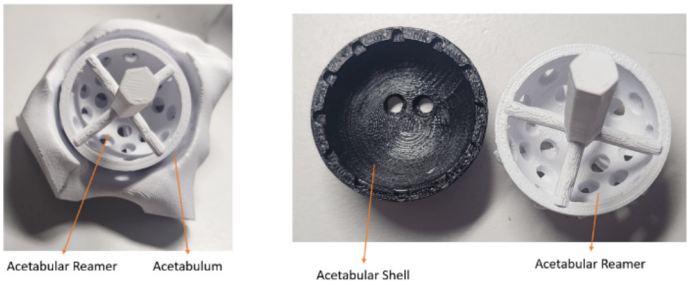


Fig. 16. 3D Printing results of acetabular reamer.

Figure 17 describes the final process of the experimental test for design optimization of the acetabular reamer as a THA implant insertion tool by attaching the reamer to a cordless screwdriver (A), visually reaming the gypsum with depth according to the shape of the acetabulum (B), comparing the depth and diameter of the acetabular shell against the reaming result (C).

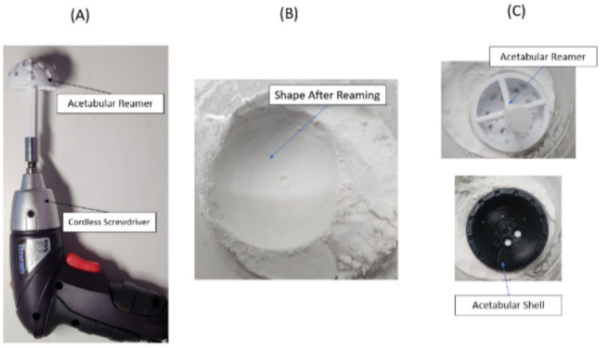


Fig. 17. Results of reaming with gypsum.

4 Discussion

To interpret the results, we have two different components, consist of femoral broach and acetabular reamer. In femoral broach, the simulation broaching cancellous bone using a femoral broach using FEA obtained maximum VMS at several parts of the broach tip. This happens because the broach contacts the cancellous bone. The VMS criteria calculate which combination of stresses at a particular point will cause failure. The higher the VMS value, the greater the risk of failure at that point or area. When the stress that occurs does not exceed the yield strength, it means that the design and materials used are safe. The femoral broach design based on this experiment concludes to be safe because the VMS value is still below the yield strength, namely 170–300 MPa.

Testing of the femoral broach design has been carried out by previous research with experimental testing on cadavers and case reports. Brzezinski et al. performed a case report regarding the unique complication of femoral broach fracture and anchorage during THA [26]. In this case report, there were 2 cases where in the process of broaching cancellous bone using a femoral broach, a fracture was found in the handle of the femoral broach [27].This is because the broach has a smaller trunnion diameter, so they modified it to a larger diameter and it does not break easily and reduces the risk of failure. However, the case report does not explain in detail what size the handle they call the "trunnion" is. In this regard, the femoral broach handle design made in this study has a diameter of 6 mm with a length of 15 mm. Based on Fig. 9, the FEA test results on the femoral broach handle are at minimum VMS so it can be said that the femoral broach handle used in this study is safe and does not have a large risk of fracture.

Based on the broaching results of experimental testing and the FEA results that were 3D printed using resin, it was found that the broaching results between experimental testing and FEA were the same. However, the FEA broaching results have rougher holes than experimental tests. This is because the materials used are different. In experimental testing, cancellous bone uses soap that is ductile so that the results of femoral broaching can match the shape of the femoral broach. Meanwhile, the FEA test uses cancellous bone material and SS316L for the femoral broach so that rough results are obtained. This is also due to the form factor of the elements and nodes used in FEA testing using Tet4 so the results obtained are less than optimal.

In acetabular reamer, according to materials engineering science, the maximum stress experienced by a material is less than its yield strength, meaning that the material has not reached its yield point or entered plastic deformation. Once the stress is removed, the material can undergo repairable deformation and return to its initial shape as it remains within the elastic deformation range [28]. However, if the stress exceeds the yield strength, it can potentially lead to permanent deformation failure of the material or structure [29]. The experimental tests help to verify the dimensional accuracy of the acetabular reamer, ensuring that the tool has dimensional cutting results following the acetabulum and acetabular shell. Based on research by Coden et al. and Macheras et al. using a reamer and implant with a larger diameter can potentially remove more bone during the reaming process, a smaller reamer and implant can provide less coverage for the implant, increasing the risk of dislocation and reduced implant stability [30, 31].

Confirming the accuracy of the reaming results, the 3D acetabular shell was placed into the perforated bar of soap, the results can be seen that the holes made matched the diameter and depth to install the acetabular shell as an implant component, and the measured fit is not too large and not too small to reduce bone removal and reduce implant dislocation. Compared with other study [32], the acetabular reamer in this study has a size that fits the acetabulum and acetabular shell, which can reduce the difficulties that previously occurred in surgical procedures. The implant placement tool is not easy to break or deform in the long term and is durable.

This research has limitations, particularly in the experimental testing using PLA and soap materials. PLA assumed cortical bone, acetabular reamer, and femoral broach. While soap and gypsum are assumed to be cancellous bone and acetabular bone. For further research, FEA testing using element and node shapes that are better than Tet4 is needed. In addition, it is necessary to manufacture femoral broach and acetabular reamer designs with SS316L material and carry out experimental testing processes on cadavers or animal femur bones that have the same characteristics as femur bones.

The biomechanical simulations and experimental results described in this paper provide significant data on the mechanical properties of the femoral broach and acetabular reamer during total hip arthroplasty. Femoral broach result suggests broaching in total hip arthroplasty causes significant local stress on the proximal femur, but does not cause excessive deformation. This structural construction of the femur is critical to ensure long-term fit, alignment, and stability, thereby minimizing the risk of complications. The acetabular reamer simulation and experimental data show that the reaming process can form the acetabular cavity with minimal stress or deformation. The stress distribution is well controlled, indicating that the risk of damage to the acetabular bone can be minimized. Overall, these findings can inform simulation and experimental results show the potential of finite element analysis in predicting biomechanical behavior during THA. This can inform the design and use of surgical tools in THA have potentially enhance patient outcomes minimizing the risk of complications related to implant placement.

5 Conclusion

This study aims to optimize the design performance of implant placement tools in patient-specific THA which may have implications for the manufacturing process. The results of this study show that the von Mises stress values of the acetabular reamer of 150.7 MPa

and femoral broach of 76.206 MPa. The maximum VMS value is lower than the yield strength value of 316L Stainless Steel material of 170–300 MPa. By this number, the designs of the acetabular reamer and femoral broach are not expected to suffer permanent deformation or failure. The experimental testing shows that the performance of the implant placement tool could work according to its function. Since the design, computer simulation, and experimental testing are compatible with each other, the future work must be manufacturing the design using biocompatible materials, such as SS316L, and implementing it clinically in a cadaveric study.

References

1. Miettinen HJA, Mäkirinne-Kallio N, Kröger H, Miettinen SSA (2021) Health-related quality of life after hip and knee arthroplasty operations. Scand J Surg 110:427–433. https://doi.org/10.1177/1457496920952232
2. Varacallo M, Luo TD, Johanson NA (2024) Total knee arthroplasty techniques
3. Fontalis A, Epinette J-A, Thaler M, Zagra L, Khanduja V, Haddad FS (2021) Advances and innovations in total hip arthroplasty. SICOT-J 7:26. https://doi.org/10.1051/sicotj/2021025
4. Supra R, Supra R, Agrawal KD (2023) Surgical approaches in total hip arthroplasty. J Orthop Sport Med. https://doi.org/10.26502/josm.511500106
5. Batailler C, Fary C, Servien E, Lustig S (2018) Influence of femoral broach shape on stem alignment using anterior approach for total hip arthroplasty: a radiologic comparative study of 3 different stems. PLoS One 13:1–11. https://doi.org/10.1371/journal.pone.0204591
6. Ferguson TA, Matta JM (2022) The anterior approach surgical technique for total hip arthroplasty: the "Matta Method" of 2020. In: Anterior hip replacement. Springer International Publishing, Cham, pp 21–38. https://doi.org/10.1007/978-3-030-91896-5_3
7. Okowinski M et al (2021) Ten-year comparison of two different techniques for femoral bone cavity preparation—broaching versus compaction in patients with cementless total hip arthroplasty. Bone Joint Open 2:1035–1042. https://doi.org/10.1302/2633-1462.212.BJO-2021-0152.R1
8. Greenhill DA, Abbasi P, Darvish K, Star AM (2017) Broach handle design changes force distribution in the femur during total hip arthroplasty. J Arthroplasty 32:2017–2022. https://doi.org/10.1016/j.arth.2016.12.031
9. Sherman WF et al (2021) Variability of cutting and thermal dynamics between new and used acetabular reamers during total hip arthroplasty. Arthroplast Today 7:91–97. https://doi.org/10.1016/j.artd.2020.12.002
10. Preutenborbeck M, Reuter J, Ferrari E (2020) Quantitative characterisation of impaction events during femoral broaching in total hip arthroplasty. Med Eng Phys 76:13–19. https://doi.org/10.1016/j.medengphy.2019.12.004
11. Nouri A, Wen C (2021) Stainless steels in orthopedics. In: Structural biomaterials. Elsevier, pp 67–101. https://doi.org/10.1016/B978-0-12-818831-6.00008-2
12. van Straten B, Dankelman J, van der Eijk A, Horeman T (2021) A circular healthcare economy; a feasibility study to reduce surgical stainless steel waste. Sustain Prod Consum 27:169–175. https://doi.org/10.1016/j.spc.2020.10.030
13. Winters G, Nutt M (2003) Stainless steels for medical and surgical applications. ASTM International, West Conshohocken, PA. https://doi.org/10.1520/STP1438-EB
14. Arunkumar K, Delhiraj U, Kumar SS (2021) Morphologic and morphometric study of human acetabulum and its clinical significance. J Clin Diagn Res. https://doi.org/10.7860/JCDR/2021/46566.14556

15. Karns MR, Patel SH, Kolaczko J, Liu RW, Mather RC, White BJ, Nho SJ, Salata MJ (2016) Acetabular rim length: an anatomical study to determine reasonable graft sizes for labral reconstruction. J Hip Preserv Surg. https://doi.org/10.1093/jhps/hnw038
16. Emara AK, Ng M, Krebs VE, Bloomfield M, Molloy RM, Piuzzi NS (2021) Femoral stem cementation in hip arthroplasty: the know-how of a "Lost" Art. Curr Rev Musculoskelet Med 14:47–59. https://doi.org/10.1007/s12178-020-09681-5
17. Karia M et al (2023) Enhancing acetabular reaming accuracy: optimal techniques and a novel reamer design. J Orthop Surg Res 18:583. https://doi.org/10.1186/s13018-023-03888-1
18. Li AJ, Lim K, Hsiung B-CB, Chen WC (2023) Stability charts for fill slopes based on finite element limit analysis methods. Soil Mech Found Eng 60:340–347. https://doi.org/10.1007/s11204-023-09899-8
19. Asmaria T, Mayasari DA, Febrananda ADG, Nurul N, Rahyussalim AJ, Kartika I (2022) Computed tomography image analysis for Indonesian total hip arthroplasty designs. Int J Electr Comput Eng 12:6123–6131. https://doi.org/10.11591/ijece.v12i6.pp6123-6131
20. Learning GD, Engineering C (2022) Synthesis and analysis of 3D shapes with geometric deep learning in computer-aided engineering
21. Bi Z, Wang X (2020) Computer aided geometric modelling. In: Computer aided design and manufacturing, pp. 35–83. https://doi.org/10.1002/9781119667889.ch2
22. Cunha Jr A, Yanik Y, Olivieri C, da Silva S (2024) Tresca versus Von Mises: Which failure criterion is more conservative in a probabilistic context? J Appl Mech. https://doi.org/10.1115/1.4063894
23. Nastri E, Todisco P (2022) Macromechanical failure criteria: elasticity, plasticity and numerical applications for the non-linear masonry modelling. Buildings. https://doi.org/10.3390/buildings12081245
24. Zahroh FF, Sugihartono I, Safitri ED (2019) Young's modulus calculation of some metals using molecular dynamics method based on the Morse potential. Comput Exp Res Mater Renew Energy 2:19. https://doi.org/10.19184/cerimre.v2i1.20557
25. Daneshpayeh S, Ashenai Ghasemi F, Ghasemi I (2022) Experimental investigation on mechanical properties of nanocomposites based on poly lactic acid/polyolefin elastomer reinforced with multi-walled carbon nanotubes, and graphene nanoplatelets. Polym Polym Compos 30:1–13. https://doi.org/10.1177/09673911211060943
26. Brzezinski A, Mascarenhas D, Simon M, Kayiaros S (2021) A unique complication of femoral broach fracture and incarceration during total hip arthroplasty. Arthroplast Today 11:49–53. https://doi.org/10.1016/j.artd.2021.07.011
27. Shinta BC (2016) Implan Ortopedi Menggunakan Software Ansys 15.0 : Pengaruh Geometri (Plate Dan development of design bone plate in orthopedic using Ansys 15.0 Software: effect of geometry plate and material)
28. Dupen B (2016) Applied strength of materials for engineering technology. Purdue Univ 10:152
29. Thompson MS, Flivik G, Juliusson R, Odgaard A, Ryd L (2004) A comparison of structural and mechanical properties in cancellous bone from the femoral head and acetabulum. Proc Inst Mech Eng Part H J Eng Med 218:425–429. https://doi.org/10.1243/0954411042632081
30. Coden G, Matzko C, Hushmendy S, Macaulay W, Hepinstall M (2021) Impact of acetabular implant design on aseptic failure in total hip arthroplasty. Arthroplast Today 7:60–68. https://doi.org/10.1016/j.artd.2020.11.017
31. Macheras GA, Koutsostathis SD, Lepetsos P, Anastasopoulos PP, Galanakos S, Papadakis SA (2014) THA following deformities due to congenital dislocation of the hip joint. HIP Int 24:29–32. https://doi.org/10.5301/hipint.5000168
32. Chen Z, Wang C, Jiang W, Tang N, Chen B (2017) A review on surgical instruments of knee arthroscopic debridement and total hip arthroplasty. Proc CIRP 65:291–298. https://doi.org/10.1016/j.procir.2017.05.001

Internet of Things and Entrepreneurship: A Bibliometric Insight into Technological Integration and Business Model Innovation

Fira Angelia Gafur and Priska Arindya Purnama(✉)

Bina Nusantara University, Jakarta, Indonesia
priska.purnama@binus.ac.id

Abstract. The intersection of the Internet of Things (IoT) and entrepreneurship presents significant opportunities for technological integration and business model innovation. Despite increasing research in this field of study, an understanding of how IoT impacts the entrepreneurial landscape comprehensively is still lacking. This study aims to fill this gap by conducting a bibliometric analysis of 2,115 secondary data publications from the Scopus database published between 2006 and 2024. This study employs VOSviewer to explore the current landscape of research at this nexus, identifying key trends, influential publications, and emerging themes. The findings indicate a significant rise trend in research publication over the past decade, with notable contributions from Germany, China, and the United States. Key themes identified include IoT-enabled business models, smart product development, and the enhancement of operational efficiency, which were derived from thematic clustering of the most frequently occurring keywords analysis. This study provides a comprehensive overview of the current state of research on IoT and entrepreneurship. It contributes to both the theoretical and practical understanding of this dynamic field. The implications of these findings suggest that practitioners such as entrepreneurs should integrate IoT more strategically to increase competitiveness and operational efficiency. The comprehensive mapping of the field also identifies implications for academics in future research directions to address the identified challenges and explore understudied areas.

Keywords: Bibliometric · Business model · Entrepreneurship · Innovation · Internet of Things · IoT · Technology

1 Introduction

Enterprises operate in a world where digital technology is omnipresent [1]. The Internet of Things (IoT) has become one of the most revolutionary technologies of the modern digital age, disrupting several industries and creating new opportunities for entrepreneurship [2]. An expanding number of formerly non-digital everyday products are adopting the IoT, impacting the nature of goods and services, leading to a transformation of the business model [1]. The transformative potential of IoT has sparked significant interest from academics and practitioners in understanding its impact on various aspects of the

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 71–80, 2026.
https://doi.org/10.1007/978-981-95-9233-3_6

business [3]. IoT allows real-time data collection and analysis by connecting physical devices to the internet, providing valuable insights that can be utilized to improve operational efficiency, optimize business processes, and develop more intelligent, responsive products and services [4].

The integration of IoT into entrepreneurship has given rise to new business models that focus on the power of technology to provide substantial additional value to customers [5]. IoT empowers entrepreneurs to create more innovative and adaptive solutions to evolving market needs by utilizing the power of real-time data to develop intelligent products and services [6]. However, there are some challenges to IoT adoption, including data security, device compatibility, and the lack of clear standards [7]. The potential and challenges in this dynamic field show that to understand how IoT impacts the entrepreneurial landscape comprehensively, a thorough analysis and examination of the existing literature is required.

Existing literature emphasizes how IoT can revolutionize a variety of sectors, including manufacturing, healthcare, agriculture, and logistics [8]. Despite the increasing amount of literature on IoT and its potential applications, a research gap exists regarding a comprehensive bibliometric analysis that addresses how IoT influences entrepreneurship at a global level. Previous studies primarily focus on certain sectors or technological advancements without providing an overarching view of how IoT drives business model innovation, particularly in relation to entrepreneurship. More contextual and thematic mapping of global literature is required to underscore the key trends and challenges that entrepreneurs encounter in adopting IoT technologies.

This study aims to map literature or publications on the Internet of Things (IoT) and entrepreneurship or business models at the global level. The aim is not only to create a comprehensive map of existing knowledge in this field but also to examine the insights found within the literature on technological integration and business model innovation. Bibliometric methods were applied in this study, utilizing secondary data from the Scopus database and using the VOSviewer tool to analyze and visualize the data [9]. This study proposes research questions, what is the mapping of IoT and entrepreneurship or business model, and how may IoT affect entrepreneurship or business model?.

2 Methods

A bibliometric approach has been applied in this study. Bibliometrics is a type of research methodology that evaluates published works, including scholarly publications quantitatively [10]. The term bibliometrics was introduced by Pritchard in 1969 [11]. The bibliometric technique is useful for evaluating the current status of a certain topic by considering several factors, including researchers, journals, academic institutions, and countries [12]. Bibliometrics enables academics to evaluate research collaboration across researchers, organizations or affiliations, and countries [13]. This study conducted a thorough literature search on a large publishing database [14]. Scopus is considered globally as a reliable source of scholarly publications by academics, hence this study has used Scopus as the primary source of data [15].

In this study, scholarly publications in the Scopus database are found and searched using keywords associated with IoT and entrepreneurship. The search terms used in data

mining were TITLE-ABS-KEY (“Internet of Things” OR “IoT”) AND TITLE-ABS-KEY (“Entrepreneurship” OR “Business Model”). Based on the search term used, 2,115 documents were discovered over the past 17 years from 2006 to 2024.

The VOSviewer (version 1.6.20) tool was used to evaluate the data that was downloaded from the Scopus database. Software frequently used for bibliometric analysis, VOSviewer, helps in the study of large volumes of research data to create networks of co-occurrence, co-citation, and collaboration [16]. Author and co-cited author, journal and co-cited journal, and cited reference co-citation networks are the main analyses carried out by the software [17]. Co-occurrence analysis provides insight into related topics that frequently appear collectively in scientific literature. This analysis counts the number of times a certain set of keywords occur together within a single document. A network connecting authors will be visualized as a result of co-authorship analysis [18].

3 Results

3.1 Annual Publication of IoT and Entrepreneurship or Business Model Publications

Figure 1 shows that studies on IoT and entrepreneurship or business models have been published every year since 2010. The line graph indicates a consistent increase in the number of documents published from 2014 to 2024 with 2024 still in progress. For the past 5 years, there have been 244 documents on IoT and entrepreneurship or business models in 2019; 278 documents in 2020; 280 documents in 2021; 288 documents in 2022; 273 documents in 2023; and 124 documents in 2024 as of July.

Less than ten documents were published annually between 2006 and 2013, which is a very low number of publications. This indicates that this topic has not been widely researched or that academic and industrial communities are not interested in it. Since 2014 there has been a moderate increase in the number of publications, with the number of published documents starting to increase gradually. This could indicate the beginning of growing interest and recognition of the importance of this topic. The number of publications peaked during the period 2019–2024, with more than 250 papers published per year.

Several important reasons have contributed to the significant increase in research publications on IoT and entrepreneurship within the last five years. The IoT has been made easier to embrace in industries like manufacturing, healthcare, and agriculture by technological advances in sensors, cloud computing, and wireless networks. In the era of industrial digitalization 4.0, an increasing number of organizations are integrating IoT into their operations due to the need for both product innovation and operational efficiency [19].

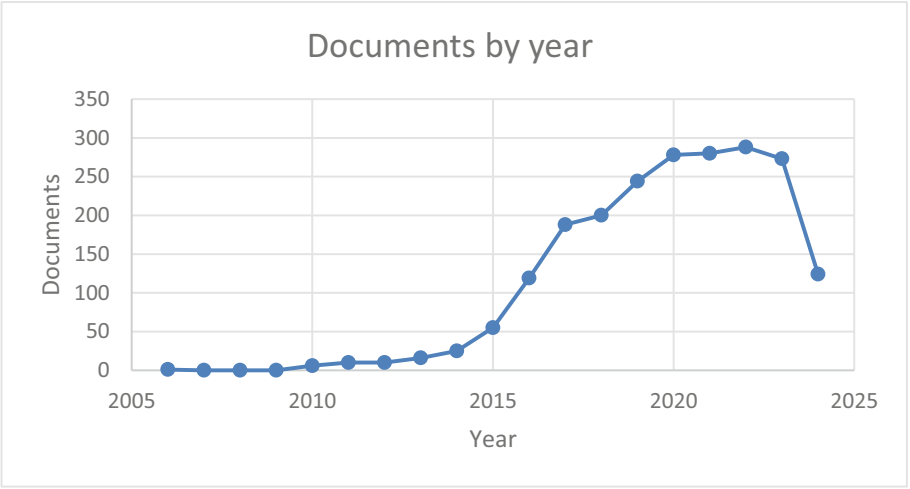


Fig. 1. Annual publications of IoT and entrepreneurship.

3.2 Documents Based on Authors of IoT and Entrepreneurship or Business Model Publications

Searching Scopus for the keywords "Internet of Things" or "IoT and "Entrepreneurship" or "Business Model" from 2006 to 2024 yields 5,541 authors who published studies on the subject. Figure 2 shows the distribution of authored documents for 10 authors. From the data shown in the graph, the most productive author is Voigt, K.I. with 13 documents published. Followed by Presser, M. and Sandkuhl, K. with 9 documents; 6 authors with the same number of publications, 7 documents; and Aagaard, A. with 6 documents.

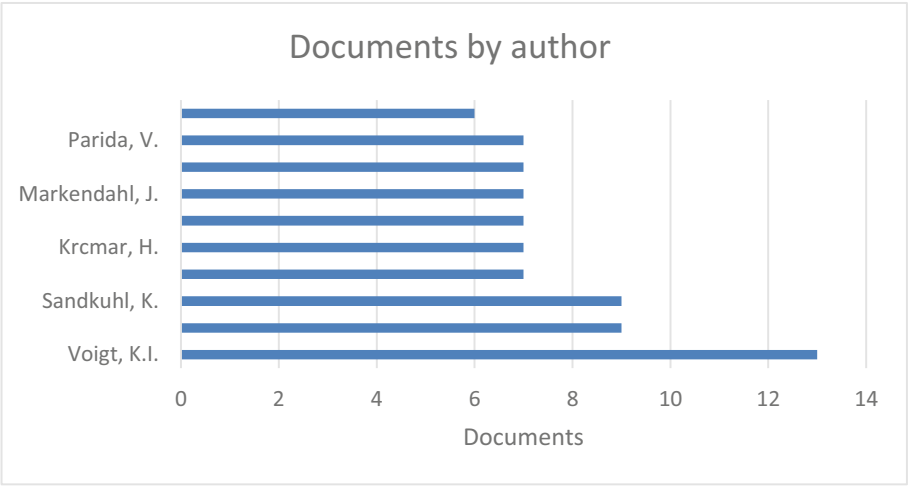


Fig. 2. Documents based on authors of IoT and entrepreneurship publications.

3.3 Documents Based on Countries of IoT and Entrepreneurship or Business Model Publications

More than 99 countries have conducted IoT and entrepreneurship research. Figure 3 shows that Germany was the country conducting the most research on IoT and entrepreneurship or business model with 286 documents in Scopus. Followed by China with 265 documents; the United States with 252 documents; India with 217 documents; the United Kingdom with 140 documents; Italy with 128 documents; Spain with 90 documents; French with 69 documents; South Korea with 63 documents; and Finland with 61 documents.

Germany, China, and the United States are three countries that demonstrate how a combination of government policy support, a developed ecosystem for research and innovation, and expertise in manufacturing and information technology [20]. The German government's backing of programs like Digital Strategy 2025 and Industrie 4.0 has driven up the use of IoT technology across a range of industrial sectors, enhancing output and efficiency [21]. China has a large electronics manufacturing base and high domestic demand for IoT products and services. The government's massive investment in technology infrastructure, including 5G networks and smart city projects, has significantly boosted the development of IoT. Major tech companies such as Xiaomi, Alibaba, and Huawei have also expanded their influence in the global IoT sector [22]. The United States boasts a highly dynamic innovation ecosystem, with Silicon Valley as its core, which drives research and development in advanced technologies and novel business models [23].

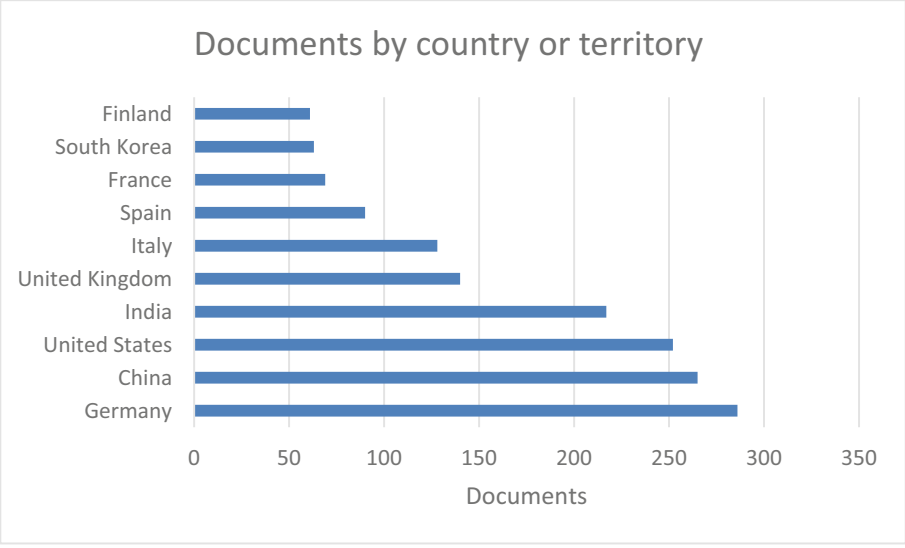


Fig. 3. Documents based on countries of IoT and entrepreneurship publications.

3.4 Theme Map of IoT and Entrepreneurship or Business Model Publications

The use of the VOSviewer tool in research allows researchers to build keyword networks and visualize study theme maps more clearly. This tool enables researchers to effectively identify emerging research areas and map connections between different topics. The IoT and entrepreneurship or business model keywords network was created for the study theme map using the VOSviewer tool. A minimum number of 5 occurrences of a keyword was required. Therefore, 682 out of 10,281 keywords were found that met the criteria. IoT and entrepreneurship is a complex and multidisciplinary field of research. Figure 4 shows there were 11 study theme groups for international academic publications on IoT and entrepreneurship or business models. These groups include IoT-enabled business models, smart product development, and the enhancement of operational efficiency. A more structured and relevant understanding of the data is provided by this thematic clustering, helping to expose underlying patterns, trends, or topics within the data.

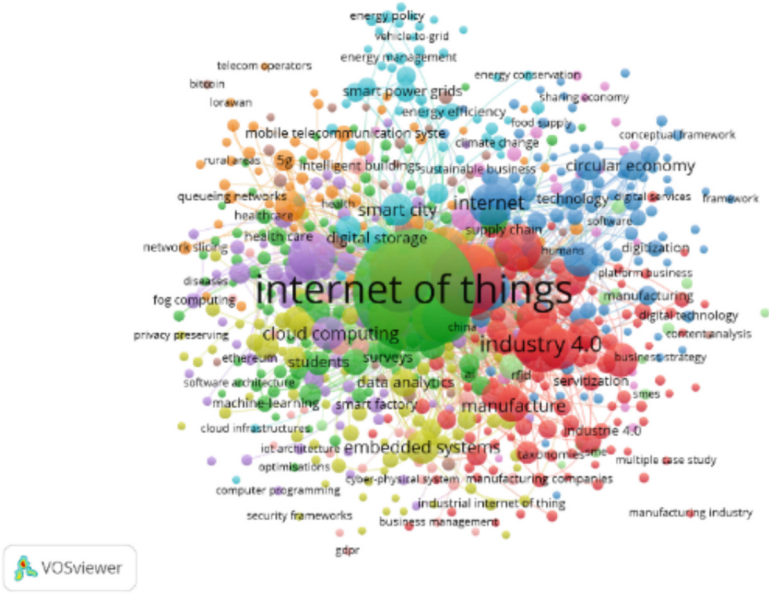


Fig. 4. Theme map of IoT and entrepreneurship publications.

3.5 Networks of Authorship

Numerous research groups have researched IoT and entrepreneurship. Based on the IoT and entrepreneurship researcher system, the VOSviewer program was used to generate the authorship network map. A minimum of six papers was required for each researcher to accumulate a certain number of research publications. As a result, only 35 out of 5,398 scholars were determined to meet the requirements. There were four community collaboration networks among international researchers in publications on IoT and entrepreneurship or business models as shown in Fig. 5. Cluster 1 (red) includes Li, H.,

Li, J., and Zhang, X.; cluster 2 (green) includes Li, Y., Wang, X., and Yu, X.; cluster 3 (blue) includes Wang, H. and Wang, J.; and cluster 4 (yellow) includes Huang, Y. and Yang, Y.

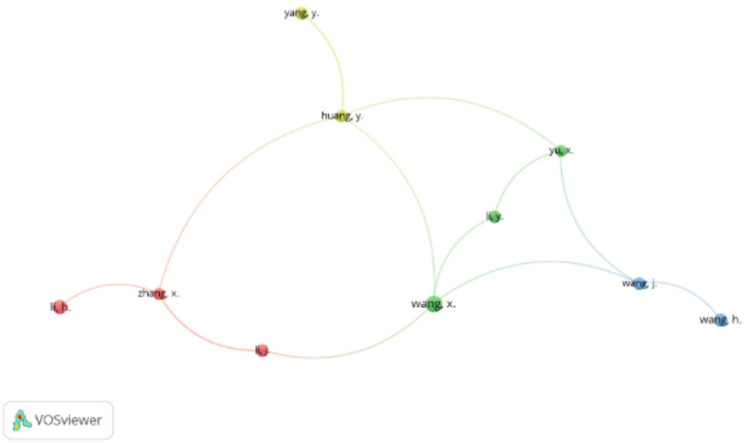


Fig. 5. Authorship network map.

4 Discussion

The findings indicate a growing body of studies examining how IoT reshaping entrepreneurship, particularly in the development of new business models and the creation of smart products [24]. The bibliometric analysis of the study demonstrates how IoT is becoming more widely acknowledged for its revolutionary potential across a range of industries, which is indicative of a larger trend of digitalization in the global economy [25]. Notable contributions from nations like Germany, China, and the United States indicate an increasing academic and practical interest in understanding IoT's impact on entrepreneurship, as reflected by the large increase in publications since 2014. These findings not only show that IoT has the potential to transform the entrepreneurship landscape globally, but also requires a deep understanding and a structured approach to address the challenges involved.

Earlier studies on comparable subjects have brought up some interesting issues of IoT and entrepreneurship. IoT and entrepreneurship were examined by Agung Purnomo, et al. (2023) in a research paper titled "Internet of Things in Entrepreneurship: The Intellectual Structure Perspective" using the same study methodology—bibliometrics [26]. Previous studies have focused more on the knowledge structure in this field without providing a deeper analysis of publication trends and contributions from the countries involved. This article fills this gap by providing a more complete and contextual analysis of the development of related literature, as well as identifying opportunities and challenges in the integration of IoT into entrepreneurship across global regions.

Comparing this study with previous bibliometric analyses on similar topics, it becomes clear that IoT's integration into entrepreneurship is gaining momentum at a faster rate than other technological trends in the business sector, such as blockchain or

artificial intelligence [27]. This suggests that IoT is perceived as having a more immediate and practical application for business innovation and efficiency. However, unlike blockchain, which has garnered significant attention for its potential to disrupt business models, IoT's impact appears to be more focused on enhancing operational efficiencies and creating smart products.

Limitations of the proposed method include its dependence on a single database (Scopus), which can restrict the scope of available research. This could lead to an incomplete picture, considering that important contributions to IoT and entrepreneurship research are also hosted by other databases, such as Web of Science and IEEE Xplore. Despite the limitations, the implications of this study expand the theoretical understanding of the impact of IoT on entrepreneurial business models and operational efficiency. Practically, the results of the study can help companies adopt IoT technology for business innovation, improve efficiency, and design more effective strategies.

5 Conclusion

Over the past decade, the Internet of Things (IoT) and entrepreneurship or business models have been increasingly popular issues among academics and business practitioners. The bibliometric analysis presented in this study provides a comprehensive insight into the development of IoT and entrepreneurship or business model publications from 2006 to 2024 as of July. The results of the bibliometric analysis indicate that the number of publications has increased moderately since 2014 and peaked between 2019 and 2024. Germany is the most productive country with 284 documents in Scopus. The most productive author is Voigt, K.I. with 13 documents published. Furthermore, the identification of the authors, countries, and theme map of the study of IoT and entrepreneurship gives an overview of how knowledge in this area is distributed globally.

The Internet of Things (IoT) has significantly impacted entrepreneurship and disrupted traditional business models by enabling new ways of creating value, optimizing operations, and enhancing customer experiences. Entrepreneurs can leverage IoT to create smart products that offer enhanced functionalities and real-time feedback to meet consumer demands more effectively. This impact is well-documented through bibliometric analysis, which provides valuable insights into research trends and the evolving landscape of IoT applications in entrepreneurship. These findings include more innovative and efficient business strategies, which companies can implement to remain competitive in the global market.

This study contributes to the literature by offering a thematic mapping that helps understand the global distribution of knowledge and identifies key trends and themes in IoT and entrepreneurship research. However, this study has limitations, including reliance on data from a single database (Scopus) and limitations in temporal coverage and analysis methodology. In addition, bibliometric analysis does not always provide in-depth insights into the quality or context of publications. Therefore, future research is recommended to explore more databases and expand the temporal coverage and methodology used to obtain a more holistic and comprehensive picture of the IoT and entrepreneurship research landscape.

References

1. Yoo Y, Boland RJ, Lyytinen K, Majchrzak A (2012) Organizing for innovation in the digitized world. Organ Sci 23:1398–1408. https://doi.org/10.1287/ORSC.1120.0771
2. Gubbi J, Buyya R, Marusic S, Palaniswami M (2013) Internet of Things (IoT): a vision, architectural elements, and future directions. Futur Gener Comput Syst 29:1645–1660. https://doi.org/10.1016/J.FUTURE.2013.01.010
3. Allioui H, Mourdi Y, Allioui H, Mourdi Y (2023) Exploring the full potentials of IoT for better financial growth and stability: a comprehensive survey. Sensors 23:8015. https://doi.org/10.3390/S23198015
4. Xu LD, He W, Li S (2014) Internet of things in industries: a survey. IEEE Trans Indus Inf 10:2233–2243. https://doi.org/10.1109/TII.2014.2300753
5. Kumar S, Tiwari P, Zymbler M (2019) Internet of Things is a revolutionary approach for future technology enhancement: a review. J Big Data 6:1–21. https://doi.org/10.1186/S40537-019-0268-2/FIGURES/9
6. Atzori L, Iera A, Morabito G (2010) The Internet of Things: a survey. Comput Netw 54:2787–2805. https://doi.org/10.1016/J.COMNET.2010.05.010
7. Alenizi AS, Al-Karawi KA (2023) Internet of Things (IoT) adoption: challenges and barriers. In: Lecture notes in networks and systems, vol 464, pp 217–229. https://doi.org/10.1007/978-981-19-2394-4_20
8. Engineering WTWA of R in S (2024) IoT and its potential for transforming industries. Int J Netw Syst 13, 27–31. https://doi.org/10.30534/IJNS/2024/041312024
9. Purnomo A, Sari YKP, Sugiantoro ADS. Veronika C (2023) Social entrepreneur study: lesson from bibliometric application during four decades. AIP Conf Proc. https://doi.org/10.1063/5.0109156/2884696
10. Purnomo A, Afia N, Rosyidah E, Maulana FI, Dina A, Sugiantoro S (2023) Mapping of E-entrepreneurship research: bibliometric review. https://doi.org/10.1051/e3sconf/202342602013
11. Glänzel W (2024) Bibliometrics as a research field: a course on theory and application of bibliometric indicators. https://www.researchgate.net/publication/242406991_Bibliometrics_as_a_research_field_A_course_on_theory_and_application_of_bibliometric_indicators. Accessed 17 Aug 2024
12. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM (2021) How to conduct a bibliometric analysis: an overview and guidelines. J Bus Res 133:285–296. https://doi.org/10.1016/J.JBUSRES.2021.04.070
13. Wube MC, Atwal H (2024) Supply chain management of micro, small, and medium enterprises (MSMEs) in Africa: a bibliometric analysis. J Innov Entrep. 13:1–13. https://doi.org/10.1186/S13731-024-00388-2/FIGURES/6
14. Purnomo A, Susanti T, Sari AK, Firdaus M, Dewi R (2020) A study of digital entrepreneurship through bibliometric visualizing from 1993 to 2019. In: Proceedings of 2020 international conference on information management and technology, ICIMTech 2020, pp 911–915. https://doi.org/10.1109/ICIMTech50083.2020.9211270
15. Purnomo A, Susanti T, Rosyidah E, Firdausi N, Idhom M (2022) Digital economy research: thirty-five years insights of retrospective review. Proc Comput Sci 197:68–75. https://doi.org/10.1016/J.PROCS.2021.12.119
16. 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/FIGURES/7
17. Li Y et al (2024) Knowledge mapping of exosomes in prostate cancer from 2003 to 2022: a bibliometric analysis. Discov Oncol 15:307. https://doi.org/10.1007/s12672-024-01183-x

18. Pan X, Yan E, Cui M, Hua W (2018) Examining the usage, citation, and diffusion patterns of bibliometric mapping software: a comparative study of three tools. J Inf 12:481–493. https://doi.org/10.1016/J.JOI.2018.03.005
19. Debnath D, Chettri SKr (2021) Internet of Things: current research, challenges, trends and applications, pp 679–694. https://doi.org/10.1007/978-981-33-4604-8_52
20. van Someren TCR, van Someren-Wang S (2013) Innovation policies in China, the USA, and the EU. Manage Professionals F402:57–112. https://doi.org/10.1007/978-3-642-36237-8_3
21. Winter J, Frey A, Biehler J (2022) Towards the next decade of Industrie 4.0—current state in research and adoption and promising development paths from a german perspective. Science. https://doi.org/10.3390/SCI4030031
22. Babel W (2022) Industry 4.0, China 2025. In: IoT (Internet of Things), pp 7–46. https://doi.org/10.1007/978-3-658-37852-3_2
23. Steiber A (2022) The Silicon valley model. Manage Profess F364:25–28. https://doi.org/10.1007/978-3-030-95754-4_3
24. Henríquez R, Xavier Martínez de Osés F, Martínez Marín JE (2020) IoT-driven business model innovation. In: Lecture notes on data engineering and communications technologies, vol 41, pp 302–314. https://doi.org/10.1007/978-3-030-34986-8_22
25. Lee I, Lee K (2015) The Internet of Things (IoT): applications, investments, and challenges for enterprises. Bus Horiz 58:431–440. https://doi.org/10.1016/J.BUSHOR.2015.03.008
26. Purnomo A, Asitah N, Rosyidah E, Pangestu G, Meiryani, Maulana FI (2023) Internet of things in entrepreneurship: the intellectual structure perspective. In: Lecture notes on data engineering and communications technologies, pp 247–257. Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-981-99-1767-9_19
27. Meyer N, Ben Said F, Alkathiri NA, Soliman M (2023) A scientometric analysis of entrepreneurial and the digital economy scholarship: state of the art and an agenda for future research. J Innov Entrep 12:1–26. https://doi.org/10.1186/S13731-023-00340-W/TABLES/9

Classification of Basic Taekwondo Kicks Using Frequency Domain Analysis and Machine Learning Algorithms

Qoriina Dwi Amalia[1], Achmad Rizal[1(✉)], Bayu Erfianto[2], and Istiqomah[1]

[1] School of Electrical Engineering, Telkom University, Bandung, Indonesia
achmadrizal@telkomuniversity.ac.id
[2] School of Computing, Telkom University, Bandung, Indonesia

Abstract. Taekwondo is one of the most popular martial arts. Several studies on martial arts have been conducted, especially Taekwondo, using sensors and cameras. One of the most vigorous research projects is the use of IMU sensors for motion and camera detection based on image processing. The use of IMU sensors and cameras is essential as both can provide accurate and real-time data on body movements, with IMUs offering details on orientation and acceleration, while image processing-based cameras provide comprehensive visualization and analysis of movements, so the combination of the two can improve accuracy and efficiency in motion detection and analysis. Taekwondo re-research is important because it can provide innovative solutions for more accurate and objective scoring in competitions, as well as contribute to improving the quality of training for athletes and practitioners, by offering efficient, self-directed, and data-driven methods of learning kick techniques, to accelerate the process of skill mastery and improve performance in competition. In this study, testing was carried out on a dataset taken from Taekwondo athletes consisting of five data classes, namely Upper Ap Chagi, Bottom Ap Chagi, Upper Dollyo Chagi, Bottom Dollyo Chagi, and Dwi Chagi, using the frequency domain. The features in the frequency domain used include Total Power (TTP), Median Frequency (MDF), 1st Spectral Moment (SM1), Peak Frequency (PKF), Mean Power (MNP), Mean Frequency (MNF), and others. With several features, testing will be carried out using several machine learning algorithms. From the test results, the highest accuracy is 97.14% using Random Forest. This result has improved considerably with the approach using the proposed frequency domain because similar tests have yet to be carried out on the dataset.

Keywords: Taekwondo · Feature extraction · Frequency domain · Classification · Machine learning

1 Introduction

The popularity of martial arts has been on the rise lately, with Taekwondo, Karate, and Muay Thai attracting the interest of many [1]. These forms of martial arts are usually practiced in groups, providing physical fitness and mental exercise for the practitioners

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 81–91, 2026.
https://doi.org/10.1007/978-981-95-9233-3_7

[2]. As the demand for martial arts training increases and so does the need for effective self-training methods, an innovation was made to create a system that could accompany athletes' training. This research is motivated by developing a companion facility supporting athlete's in-home training. Such a system would be helpful for athletes preparing for belt exams. Mastery of basic techniques, such as poomsae in Taekwondo, is crucial during the exam [3]. With independent practice, athletes can improve the quality of the kicks learned.

Previous research used statistical feature extraction for Taekwondo kick classification. This research has achieved an accuracy of 99.35% [3]. Due to the controller's limitation for implementation, the statistical feature was directly calculated from the acquired data [4]. Another study with a different dataset containing two classes of Taekwondo kicks with Fast Fourier Transform (FFT) feature extraction. The resulting accuracy was 91.6% for the Support Vector Machine (SVM) and 96.8% with the K-Nearest Neighbor (KNN) [5]. Another research study raised the same topic, focusing on model implementation for the Raspberry Pi Zero W controller [4]. The first study utilizing statistical features may not fully capture important spectral information in frequency signals. Meanwhile, the second study utilizing the frequency domain tends to have limitations in terms of generalization and the ability to handle more complex classification problems, mainly due to the limited number of classes and the size of the controller components used, which may hinder the movement of athletes.

FFT is used to transform signals from time to frequency domain. This method is often used for biomedical signal analysis with impressive results. The use of FFT for ECG signal analysis is reported in [6]. Meanwhile, a related study used FFT to compute frequency domain features for heart sound analysis [7]. Meanwhile, other researchers used a combination of heart sound characteristics in the time and frequency domains [8]. Features in the frequency domain are also utilized for EEG signal analysis [9, 10].

Seeing the potential of using FFT for signal analysis, the research proposed an approach to frequency domain analysis for kick classification in Taekwondo. Classification is performed on a dataset consisting of five classes. This approach has not been made much, considering this dataset is a primary dataset taken on five subjects in black belt Taekwondo athletes. This research will provide a different perspective as the frequency domain can help separate relevant signals from noise, resulting in cleaner data that is easier to analyze. This research can certainly overcome the shortcomings of previous research because it uses the frequency domain and the same number of classes, five classes. Additionally, the frequency domain allows machine learning models to more effectively recognize repetitive and consistent patterns in kicks, which may not be detected through simple statistical analysis. The frequency domain representation can also provide a more precise visualization of the signal characteristics to aid in data analysis and interpretation. Therefore, the proposed model is expected to improve the accuracy and efficiency of the kick classification system in Taekwondo using machine learning.

2 Materials and Method

This section explains the process from dataset processing to classification. Figure 1 presents the flow of the entire research process.

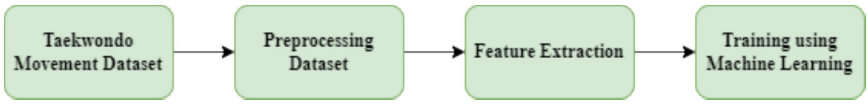


Fig. 1. Proposed method

2.1 Dataset

The dataset used in this study is a primary dataset derived from Taekwondo athletes as seen in Fig. 2. Dataset retrieval is done using an Arduino Nano 33 BLE device equipped with an IMU sensor [11]. The sensor is hooked on the athlete's ankle. The selection of athletes is based on their belt level. Athletes with black belts are considered to have mastered movement techniques better than other belt levels [12]. This dataset consists of five classes: Upper Ap Chagi, Bottom Ap Chagi, Upper Dollyo Chagi, Bottom Dollyo Chagi, and Dwi Chagi [3, 13, 14].

Fig. 2. Procedure for collecting information

2.2 Feature Extraction

The Fast Fourier Transform (FFT) converts a signal from the time domain to the frequency domain, which identifies the main frequency components of the signal [15]. This can be useful in recognizing motion patterns from sensor data, such as accelerometers [16]. The FFT can be defined as follows [Eq. (1)].

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}kn} \tag{1}$$

where $X[k]$ is the frequency component at k, $x[n]$ is the signal value at time n, j is an imaginary number $\sqrt{-1}$, N is the number of samples in the signal, and k is the frequency index from 0 to $N-1$. To normalize the FFT result, it is often divided by N [15].

Table 1 shows the frequency domain features used in this study. The frequency domain will show a different pattern compared to using statistics alone. The domain

frequency features used include Median Frequency (MDF), Mean Power (MNP), Total Power (TTP), 1st Spectral Moment (SM1), 2nd Spectral Moment (SM2), 3rd Spectral Moment (SM3), Frequency Ratio (FR), Power Spectrum Ratio (PSR), Mean Frequency (MNF), Peak Frequency (PKF), and Variance of Central Frequency (VCF).

Table 1. Frequency domain features used

No.	Name/equation	Definition
1	Mean frequency $MNF = \sum_{j=1}^{M} f_y P_j / \sum_{j=1}^{M} P_j$	The average frequency is the sum of the product of the EMG power spectrum and the quotient of the frequency divided by the sum of the spectrum intensities
2	Median frequency $\sum_{j=1}^{MDF} P_j = \sum_{j=MDF}^{M} P_j = \frac{1}{2} \sum_{j=1}^{M} P_j$	The frequency at which the spectrum is divided into two regions of equal amplitude
3	Peak frequency $PKF = \max(P_j), j = 1, \ldots, M$	Existing frequencies based on the maximum power that occurs
4	Mean power $MNP = \sum_{j=1}^{M} P_j / M$	Average power of the EMG power spectrum
5	Total power $TTP = \sum_{j=1}^{M} P_j = SM0$	The sum or entirety of the power spectrum of the EMG
6	1st spectral moment $SM1 = \sum_{j=1}^{M} P_j f_j$	The power spectrum of signal can also be extracted statistically using spectral moments, of which SM1–SM2 are the most significant
7	2nd spectral moment $SM2 = \sum_{j=1}^{M} P_j f_j^2$	
8	3rd spectral moment $SM3 = \sum_{j=1}^{M} P_j f_j^3$	
9	Frequency ratio $FR = \sum_{j=LLC}^{ULC} P_j / \sum_{j=LHC}^{UHC} P_j$	Aims to distinguish between muscle contraction and relaxation, using the ratio between low frequency and high frequency

(*continued*)

Table 1. (*continued*)

No.	Name/equation	Definition
10	Power spectrum ratio $PSR = \frac{P_0}{P} = \sum_{j=f_0-n}^{f_0+n} P_j / \sum_{j=-\infty}^{\infty} P_j$	The ratio between the energy P0 found near the maximum value of the EMG power spectrum and P is the overall energy of the EMG power spectrum
11	Variance of central frequency $VCF = \frac{1}{SM0} = \sum_{j=1}^{M} P_j(f_j - f_c)^2 = \frac{SM2}{SM0} = \left(\frac{SM1}{SM1}\right)^2$	Key characteristics of PSD

In addition to using the above features, in the test several statistical features are added which are measured through FFT, the features used include Mean, Median, Max, Min, Skewness, Kurtosis, Variance, and Entropy [17].

2.3 Classifier

Some of the methods used in this research include KNN, SVM, RF, DCT, and GNB [18]. The following is an explanation of each classifier used.

K-Nearest Neighbor. For classification and regression applications, KNN is a non-parametric learning method based on examples [19]. Using a distance measure, the procedure is locating the k closest data points to a new data point, then allocating the new point to the most prevalent class or value among those neighbors [20].

Support Vector Machine. SVM is a learning method that's applied to regression analysis and classification. The process entails determining which hyperplane is best for maximizing the margin between data points or classes. Creating a decision boundary that divides classes by the greatest distance from the closest data points is the aim [21, 22].

Random Forest. The Random Forest ensemble learning technique was used by many decision trees to decrease overfitting and improve prediction accuracy. Using randomly chosen subsets of training data and characteristics; different decision trees are created. Their predictions are then combined using majority voting for classification or averaging for regression [23, 24].

Decision Tree. A tree-based learning technique called Decision Tree is employed for problems involving regression and classification. Recursive partitioning of the data according to feature values that maximize information gain or minimize impurity is the method used in this process. Decision nodes, which represent feature tests, and leaf nodes, which indicate class labels or predicted values, make up the resultant tree structure [25].

Gaussian Naive Bayes. Gaussian Naive Based on the independence of characteristics and the Bayes theorem, the Bayes classifier is probabilistic. The procedure entails calculating the likelihood that each class will have a given feature value, after which the class with the highest likelihood is assigned to a new data point. The Gaussian distribution of the characteristics is assumed by the algorithm [26, 27].

3 Result

3.1 Dataset

Data collection on Taekwondo athletes contains acceleration values in the form of axis values from the x, y, and z axes (Fig. 3). Each class includes 14 data, with a total of 70 data. Figure 3 shows data from one of the existing classes, the Upper Ap Chagi class [16, 26]. The horizontal axis shows time in specific units, while the vertical axis shows acceleration values measured in standard units. Figure 3 is data from one of the existing classes [28].

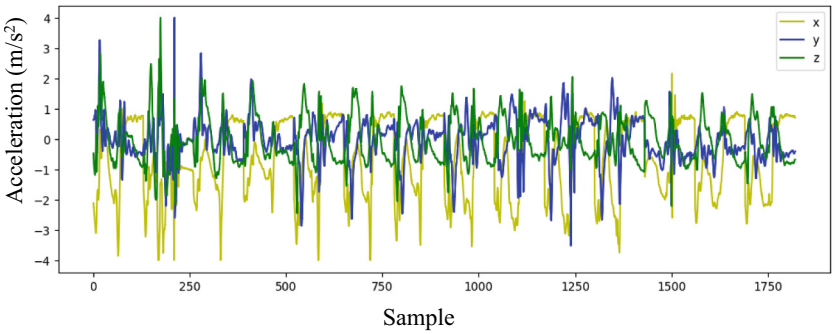


Fig. 3. Acceleration of Upper Ap Chagi class data

3.2 Fast Fourier Transform (FFT)

Figure 4 shows the frequency spectrum generated from the FFT analysis for the Bottom Dollyo Chagi movement [29]. The graph displays the distribution of signal energy at various frequencies, with frequency on the horizontal axis measured in Hertz (Hz) and amplitude on the vertical axis. From the graph, low frequencies dominate with higher amplitudes, indicating that the low-frequency component has a significant influence on the acceleration pattern of this movement. Frequency spectrum analysis like this is crucial in understanding the dynamic characteristics of an athlete's movement, as it allows researchers to identify the key frequencies that contribute to the effectiveness and smoothness of the movement [6].

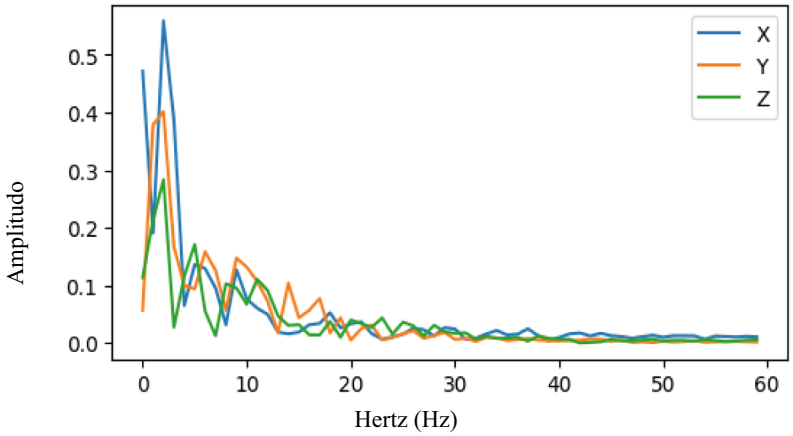


Fig. 4. Frequency spectrum of FFT result for Bottom Dollyo Chagi

3.3 Accuracy Results

Based on the existing research scenario, this study begins with data collection on Taekwondo athletes. The data obtained will then be processed first through the pre-processing stage to ensure that it is in an optimal condition for further analysis. After the pre-processing stage is complete, the data will be analyzed using the method proposed in this study. The results of this analysis will then be evaluated through the classification process using five different types of classifiers to assess the performance of each classifier against the features extracted from the frequency domain.

Table 2. Accuracy results using frequency domain feature

Classifier				
KNN (%)	SVM (%)	RF (%)	DCT (%)	GNB (%)
78.57	64.29	97.14	85.71	71.43

Table 2 presents the accuracy results of the classification process conducted using five classifiers, namely KNN, SVM, RF, DCT, and GNB. The classification results show performance variations among the five classifiers, with Random Forest (RF) showing the highest accuracy of 97.14%, followed by DCT with an accuracy of 85.71%. KNN also performed reasonably well with an accuracy of 78.57%, while SVM and GNB recorded lower accuracies of 64.29% and 71.43% respectively. These results suggest that more complex classification methods, such as Random Forest, can utilize frequency domain features more effectively, providing higher accuracy in classifying Taekwondo athletes' movement data.

4 Discussion

Our study shows that the Random Forest algorithm is very good at finding complex patterns in data when using all the frequency domain features. It achieved an impressive accuracy of 97.14%, which is better than the other algorithms tested. The DCT classifier also did well, with an accuracy of 85.71%, showing it can handle frequency domain data effectively. However, the KNN and GNB classifiers had lower accuracies of 78.57% and 71.43%, respectively. The SVM classifier performed the worst with an accuracy of 64.29%, suggesting it may not be the best choice for certain types of data [30].

These results align with previous research, such as studies that used the FFT for feature extraction, which reported a high accuracy of 96.8% with the KNN algorithm [5]. Although this approach demonstrated good performance, it relied heavily on statistical features without incorporating additional frequency domain features, which limited its ability to capture more complex patterns in the data. Another study using statistical features with the same dataset achieved an even higher accuracy of 99.35% [3]. However, this study also faced limitations, such as potential overfitting due to the use of a smaller feature set and the inability to explore the benefits of ensemble methods fully.

Our current research addresses these shortcomings by integrating a broader set of frequency domain features, enhancing the ability to capture intricate patterns that single-feature approaches might miss. By using the Random Forest algorithm, which employs ensemble learning techniques, we mitigate overfitting and improve the model's generalization capability. This approach not only leverages the strengths of frequency domain features but also optimizes the balance between feature complexity and model performance, leading to a more robust and accurate classification. Thus, our study provides a more comprehensive solution by addressing the limitations of feature selection and algorithm choice evident in previous research, ultimately demonstrating the superior performance of ensemble methods like Random Forest for complex datasets.

Our study has some limitations. We used a specific set of frequency domain features, which may affect how generalizable the results are. The performance of the classifiers might change with different features or datasets. Also, Random Forests are more complex and take longer to train compared to simpler models like KNN. We did not investigate how feature scaling or normalization might affect the results.

These limitations highlight the need to use feature selection and dimensionality reduction techniques to improve performance and reduce computational demands. Future research should explore how different feature sets, scaling, and normalization affect classifier performance. Trying out other advanced algorithms and ensemble methods could also offer better ways to handle complex datasets. Researchers should carefully consider the specifics of their data when choosing algorithms.

To build on our findings, future studies should investigate how different feature sets and preprocessing techniques like scaling and normalization affect classifier performance. Exploring other advanced algorithms and ensemble methods could be valuable. Additionally, examining how frequency domain features can be used in different areas, such as audio analysis and medical signal processing, could be useful.

5 Conclusion

This research aimed to assist Taekwondo athletes in practicing independently at home or supplementing their training before belt promotion exams. By using 11 features from the frequency domain and statistical features from FFT, the Random Forest algorithm achieved an accuracy of 97.14%, significantly outperforming the KNN algorithm, which only achieved 78.57%. This highlights that the Random Forest algorithm is better suited for handling complex datasets with numerous frequency domain features, unlike KNN, which struggles with high feature complexity.

These findings have significant implications for the broader field of sports training. The integration of frequency domain features with ensemble methods like Random Forest has been shown to effectively enhance the accuracy of classification tasks, particularly for Taekwondo athletes. This approach could potentially influence the field by encouraging the adoption of similar methods for other types of physical training and motion analysis, thereby improving the overall quality of sports training.

This study makes a unique contribution to the field by introducing a novel approach to processing motion datasets for Taekwondo. It focuses on underutilized frequency domain features and ensemble algorithms, providing new insights into dataset handling that have not been explored in previous research. This fills a significant gap in the literature regarding self-training applications for martial arts, making it a valuable addition to the field.

Future research should expand on these findings by collecting larger datasets and exploring deep learning techniques that further improve accuracy. Additionally, implementing these methods on smaller, portable devices could make the training process more convenient for athletes, allowing them to train without distractions.

In conclusion, this research significantly advances the field of self-training for Taekwondo athletes by demonstrating the potential of frequency domain features combined with ensemble methods like Random Forest. This approach not only improves classification accuracy but also sets a foundation for future innovations in sports training technology, emphasizing the importance of choosing the right algorithms and features for specific tasks.

References

1. Origua RS, Marks J, Estevan I, Barnett LM (2018) Health benefits of hard martial arts in adults: a systematic review. J Sports Sci 36:1614–1622. https://doi.org/10.1080/02640414.2017.1406297
2. Fachrezzy F, Maslikah U, Safadilla E, Reginald R, Hendarto S (2021) Physical fitness of the poomsae taekwondo athletes in terms of agility, balance and endurance. Kinestetik J Ilmiah Pendidikan Jasmani 5:111–119. https://doi.org/10.33369/jk.v5i1.14364
3. Amalia QD, Gunawan AA, Yulian GS, Rizal A, Istiqomah I (2023) Automated Taekwondo kick classification using SVM and IMU sensor on Arduino Nano 33 BLE. In: International conference on electronics, biomedical engineering and health informatics (ICEBEHI). https://doi.org/10.1007/978-981-97-1463-6_3
4. Jaya IGP, Dharmmesta RA, Rizal A, Istiqomah I (2022) Application foot kick classification in Taekwondo with inertia sensor and machine learning. In: 2022 IEEE Asia Pacific conference

on wireless and mobile (APWiMob), pp 1–6. IEEE. https://doi.org/10.1109/APWiMob56856.2022.10014067
5. Dharmmesta RA, Jaya IGP, Rizal A, Istiqomah (2022) Classification of foot kicks in Taekwondo using SVM (support vector machine) and KNN (K-nearest neighbors) algorithms. In: IEEE international conference on industry 4.0, artificial intelligence, and communications technology (IAICT). https://doi.org/10.1109/IAICT55358.2022.9887475
6. Fathail I, Bhagile VD (2022) ECG paper digitization and R peaks detection using FFT. Comput Intell Soft Comput. https://doi.org/10.1155/2022/1238864
7. Istiqomah, Rizal A, Chiueh H (2023) Heart abnormality classification with power spectrum feature and machine learning. In: Lecture notes in electrical engineering. Springer Science and Business Media Deutschland GmbH, pp 307–319. https://doi.org/10.1007/978-981-99-0248-4_22
8. Kristomo D, Hidayat R, Soesanti I, Kusjani A (2016) Heart sound feature extraction and classification using autoregressive power spectral density (AR-PSD) and statistics features. In: AIP conference proceedings. American Institute of Physics Inc. https://doi.org/10.1063/1.4958525
9. Al-Fahoum AS, Al-Fraihat AA (2014) Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains. ISRN Neurosci. 2014:1–7. https://doi.org/10.1155/2014/730218
10. Li H, Huang G, Lin Q, Zhao J, Fu Q, Li L, Mao Y, Wei X, Yang W, Wang B, Zhang Z, Huang D (2020) EEG changes in time and time-frequency domain during movement preparation and execution in stroke patients. Front Neurosci. https://doi.org/10.3389/fnins.2020.00827
11. Liu Z, Yang M, Li K, Qin X (2024) Recognition of TaeKwonDo kicking techniques based on accelerometer sensors. Heliyon 10:e32475. https://doi.org/10.1016/j.heliyon.2024.e32475
12. Sousa JL, Lein K, Jimenez JM, Gamonales JM (2021) Taekwondo: Korean unarmed martial art as the origin of a modern combat sport. A literature review. Congreso Int Innov Docente. https://doi.org/10.13140/RG.2.2.25923.68641
13. Arun Faisal I, Waluyo Purboyo T, Siswo Raharjo Ansori A (2019) A review of accelerometer sensor and gyroscope sensor in IMU sensors on motion capture. J Eng Appl Sci 15:826–829. https://doi.org/10.36478/jeasci.2020.826.829
14. Ishac K, Eager D (2021) Evaluating martial arts punching kinematics using a vision and inertial sensing system. Sensors 21:1–25. https://doi.org/10.3390/s21061948
15. Ambaye GA (2020) Time and frequency domain analysis of signals: a review. Int J Eng Res Technol (IJERT)
16. Cho D, Lee H, Lee W, Kang S (2021) Detecting anomalous kicks in taekwondo with spatial and temporal features. IEEE Access. 9:164928–164934. https://doi.org/10.1109/ACCESS.2021.3134967
17. Mass Duki MS, Shapiee MNA, Abdullah MA, Khairuddin IM, Mohd Razman MA, Abdul Majeed PPA (2021) The classification of Taekwondo kicks via machine learning: a feature selection investigation. Mekatronika 3:61–67. https://doi.org/10.15282/mekatronika.v3i1.7153
18. Saha SS, Sandha SS, Srivastava M (2022) Machine learning for microcontroller-class hardware: a review. https://doi.org/10.1109/JSEN.2022.3210773
19. Alodia Yusuf SA, Hidayat R (2019) MFCC feature extraction and KNN classification in ECG signals. In: International conference on information technology, computer and electrical engineering (ICITACEE). IEEE, pp 1–5. https://doi.org/10.1109/ICITACEE.2019.8904285
20. Aldi F, Nozomi I, Soeheri S (2022) Comparison of drug type classification performance using KNN algorithm. SinkrOn 7:1028–1034. https://doi.org/10.33395/sinkron.v7i3.11487
21. Gouri MS, Vijulagrace KS (2022) Comparison of SVM and KNN classifiers on an EEG signal with a simple dataset. In: Advances in parallel computing. IOS Press BV, pp 161–168. https://doi.org/10.3233/APC220022

22. Dwi N, Safitri N, Andhyka Kusuma W, Minarno AE (2021) Human activity recognition Menggunakan algoritma SVM with gridsearch. Repositor 3:433–440
23. Gröner G (2019) A Random Forest Based Classifier for Error Prediction of Highly Individualized Products. Presented at the. https://doi.org/10.1007/978-3-662-58485-9_4
24. Sheykhmousa M, Mahdianpari M, Ghanbari H, Mohammadimanesh F, Ghamisi P, Homayouni S (2020) Support vector machine versus random forest for remote sensing image classification: a meta-analysis and systematic review. IEEE J Select Topics Appl Earth Observ Remote Sens. https://doi.org/10.1109/JSTARS.2020.3026724
25. Charbuty B, Abdulazeez Mohsin A (2021) Classification based on decision tree algorithm for machine learning. J Appl Sci Technol Trends 2:20–28. https://doi.org/10.38094/jastt20165
26. Bi Z, Han Y, Huang C, Wang M (2019) Gaussian Naive Bayesian data classification model based on clustering algorithm. In: Proceedings of the 2019 international conference on modeling, analysis, simulation technologies and applications (MASTA 2019). Atlantis Press, Paris. https://doi.org/10.2991/masta-19.2019.67
27. Saravanan R, Sujatha P. A state of art techniques on machine learning algorithms: a perspective of supervised learning approaches in data classification
28. Shakya SR, Zhang C, Zhou Z (2018) Comparative study of machine learning and deep learning architecture for human activity recognition using accelerometer data. Int J Mach Learn Comput 8:577–582. https://doi.org/10.18178/ijmlc.2018.8.6.748
29. Mihaylov D, Petkov I (2022) Most used taekwondo kicking techniques. In: International scientific congress. Scientific Publishing House NSA Press, pp 240–243. https://doi.org/10.37393/ICASS2022/42
30. Hoang TT, Ahn H (2023) An evaluation method of Taekwondo Poomsae performance. J Inf Commun Converg Eng 21:337–345. https://doi.org/10.56977/jicce.2023.21.4.337

Evaluating Ensemble Learning and Neural Network Models Towards Covid-19 Detection Through Cough Recordings

William Axel Cuangdinata[1(✉)], Kent Alber Fredson[1], Alfi Yusrotis Zakiyyah[2], and Meiliana[1]

[1] Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia
william.cuangdinata@binus.ac.id

[2] Mathematics Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia

Abstract. The COVID-19 pandemic has resulted in over 520 million confirmed cases and more than 6 million deaths as of May 2022. Despite advancements in vaccine development, there is still a critical need for effective and accessible diagnostic methods. One promising area of research is analyzing cough sounds, which has shown a strong correlation with COVID-19. This study aims to offer insights to help researchers and practitioners choose the most suitable machine learning models, enhancing their effectiveness and efficiency across various fields. This study contributes to the development of more effective applications by guiding model selection. This involves pre-processing audio data, extracting pertinent features, and training various models. Using datasets such as COUGHVID and COVID-19 Recordings, the research highlights the challenges posed by dataset imbalances on model accuracy, as only 72–76% of accuracy in large dataset with some models achieving only 2–3% of F1 Score in COVID label, and 88–96% of accuracy in small dataset with 62% in F1-Score of COVID label. Despite these challenges, the evaluation provides valuable insights into the strengths and limitations of different machine learning models in the context of COVID-19 detection. This study provides a comparative analysis of machine learning models for COVID-19 detection using cough sounds, presenting crucial outcomes that guide the selection of effective AI-driven diagnostic tools for healthcare applications.

Keywords: COVID-19 · Imbalanced datasets · Convolutional neural networks · Ensemble learning · Cough analysis

1 Introduction

COVID-19, caused by the SARS-CoV-2 virus, was declared a global pandemic by the World Health Organization (WHO) in early 2020. By early 2021, over 108 million confirmed cases and nearly 2.4 million deaths had been reported globally [1]. By May 2022, confirmed cases had surpassed 520 million, reaching 524 million by late 2022,

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 92–108, 2026.
https://doi.org/10.1007/978-981-95-9233-3_8

with deaths exceeding 6 million [2, 3]. Following that, the WHO reported that there had been over 776 million confirmed cases worldwide by the end of September 2024 [4]. Various approaches have been proposed to enhance the efficacy of detecting COVID-19 in individuals, including real-time reverse transcription polymerase chain reaction (RT-PCR) [5], chest X-rays [6], chest computed tomography (CT) scans [7], as well as heart rate, sleep patterns, and activity data collected via wearable sensors [8]. Notably, studies have shown that cough analysis correlates highly with COVID-19 infection [9].

Several artificial intelligence (AI) models have been developed to identify patterns in cough audio and classify them accordingly. Among these, a combined ResNet-50 and Deep Neural Network (DNN) model achieved an impressive accuracy of 96% [9]. Other models include a linear-based Support Vector Machine (SVM) with a Radial Basis Function (RBF), which achieved an Area Under the Curve (AUC) of 56% [10], and a model combining Hidden Markov Models, Gaussian Mixture Models, and Mel Frequency Cepstral Coefficients (HMMs-GMMs-MFCC), which reached 91.57% accuracy [11]. Additionally, a Random Forest model demonstrated an accuracy of around 90% [1]. Convolutional Neural Networks (CNNs) have also shown promise in this area. Notably, AlexNet achieved an AUC of 93% [12], while VGG19 and DenseNet-201 outperformed other CNN architectures [13]. As the field of machine learning continues to evolve rapidly, selecting the most effective model for a specific application remains crucial.

While various diagnostic methods, including real-time reverse transcription polymerase chain reaction (RT-PCR), Chest X-ray, and Chest CT images, have been developed and widely used to detect COVID-19, there remains a significant gap in leveraging non-invasive, easily accessible diagnostic tools while that can be employed on a large scale. Additionally, while several machine learning models, such as ResNet-50, DNN, SVM, and Random Forest, have been applied to COVID-19 detection, there is a lack of comprehensive studies comparing the performance of different machine learning approaches, particularly neural networks against ensemble learning models, in the context of audio-based COVID-19 detection. Most studies focus on the accuracy of individual models without thoroughly examining their comparative strengths and limitations in specific tasks.

To address the existing gap in research, we aim to conduct a comprehensive comparative analysis between Neural Network models and ensemble learning models, particularly focusing on their performance across various tasks. While the accuracy of individual models has been extensively studied, there remains a significant need to understand the strengths and limitations of these models when applied to audio-based COVID-19 detection. Our study centers on two well-regarded ensemble learning models: Random Forest and XGBoost. Random Forest, known for its robustness in handling large datasets, excels in both classification and regression tasks. It has demonstrated superior performance compared to Support Vector Machines (SVM) in various studies [14]. Given its versatility and efficacy, Random Forest serves as a critical benchmark in our evaluation. We also selected XGBoost as a benchmark model due to its proven advantages over both SVM and K-Nearest Neighbors (KNN). The term ‘gradient’ in XGBoost refers to the use of the gradient descent algorithm to minimize errors, while ‘extreme’ highlights the model’s efficiency in maximizing learning outcomes and minimizing computing

resources. These features make XGBoost a powerful tool in machine learning, particularly for tasks that demand high precision [15]. By comparing these ensemble learning models against Convolutional Neural Networks (CNNs), we aim to provide valuable insights that will assist researchers and practitioners in making informed decisions about model selection. Our goal is to contribute to more effective and efficient applications in diverse fields by elucidating the circumstances under which certain models outperform others. This study not only fills a gap in the literature but also offers practical guidance for optimizing model performance based on the specific needs of different applications.

2 Materials and Method

In this section, we will summarize our research regarding COVID-19 Detection using cough sounds.

2.1 CNN Models

In [16], the authors use CNN, VGG16, VGG19, and Inception V3 models, and found that VGG19 gave the best performance. Its straightforward architecture, using consistent 3×3 convolutions and 2×2 max-pooling layers, makes it easy to implement and understand, while its depth allows for robust feature extraction. Additionally, VGG19 is widely used for transfer learning, leveraging pretrained models to reduce the need for large and labeled datasets. However, it has significant drawbacks, including high computational and memory costs, slow training and inference times, and potential redundancy in parameters. The model's considerable number of parameters, approximately 144 million, requires substantial resources, which can be limited for real-time applications or deployment on devices with limited memory.

The authors in [17] have shown that fully connected convolutional neural networks are effective at diagnosing COVID-19 infections using X-ray images. Inspired by their success, we want to test the use of fully connected neural networks to predict COVID-19 based on cough sounds. Our goal is to see if these neural networks can identify COVID-19 by analyzing the unique features in the sound of a cough.

Study [12] compares various CNN models such as VGG-16, VGG-19, LeNet-5, AlexNet, ResNet-50, and ResNet-152, and finds that AlexNet outperforms the others. Furthermore, AlexNet is a pioneering deep learning model which sets new standards for image recognition by introducing ReLU activation and dropout regularization. However, it has high computational and memory requirements, a complex architecture, and can be overfitted without proper regularization [18].

2.2 Ensemble Learning Models

When combined with the BiGRU-Attention model, XGBoost achieves state-of-the-art performance and allows for enhanced accuracy in monitoring breathing waveforms [19]. However, measuring breathing patterns in noisy environments remains challenging. Despite its complexity and high memory usage, XGBoost is known for its high performance, regularization, and scalability, making it popular for many applications. In

respiratory sound classification, its integration with advanced feature extraction techniques has proven highly effective, demonstrating its potential in medical diagnostics and other fields requiring robust classification models [20].

Whilst study [1] states Random Forest outperforms other models such as SVM, Linear Discriminant Analysis (LDA), Logistic Regression (LR), and Naïve Bayes (NB), study in [21] also suggests that the AUC of the Random Forest model in that paper is still one of the highest amongst others. By understanding both studies, we can acknowledge that Random Forest is a robust and versatile machine learning algorithm known for its high accuracy, ability to handle missing data, and resistance to overfitting. It is effective in both classification and regression tasks and provides valuable insights into feature importance. However, its drawbacks include computational intensity, longer training times, and memory usage.

2.3 MFCC Features

Mel-Frequency Cepstral Coefficients (MFCC) is widely regarded as one of the most important features in speech and audio signal processing. They are commonly used in speech recognition and audio classification. The use of MFCCs is based on the ability of these coefficients to model the shape of the vocal tract, which is reflected in the envelope of the short-term power spectrum of the sound signal. This approach is particularly useful for analyzing cough sounds, as it captures subtle variations and characteristics that may distinguish healthy individuals from those infected with COVID-19 [22].

MFCCs are derived by taking into account human auditory perception, which is more sensitive to certain frequency ranges. This is modeled using a nonlinear scale known as the Mel scale, which is more representative of how humans perceive sound frequencies. The process of computing MFCC features typically involves the following key steps [22]:

1. **Framing the Audio Signal**: The input sound signal, such as a cough, is divided into short overlapping frames (typically 20–40ms in length) to capture local temporal and spectral features of the sound. The use of overlapping frames helps preserve the temporal continuity of the audio signal.
2. **Power Spectrum Calculation**: For each frame, a Fourier transform is applied to convert the time-domain signal into its frequency components. The result is the power spectrum, which shows how the energy of the signal is distributed across different frequency bands.
3. **Mel Filter Bank**: The power spectrum is then passed through a series of triangular filters, spaced along the Mel frequency scale. This step emphasizes certain frequency ranges in a manner that mimics human auditory sensitivity, giving more importance to the lower frequencies and gradually reducing emphasis on higher frequencies.
4. **Logarithm of Filter Bank Energies**: To compress the dynamic range and make the data more manageable, the logarithm of the filter bank outputs is computed. This step helps in reducing the influence of variations in amplitude across different cough samples.
5. **Discrete Cosine Transform (DCT)**: The logarithmic Mel-scaled filter bank outputs are then subjected to a Discrete Cosine Transform (DCT). The DCT helps in decorrelating the coefficients and transforming the data into a compressed form. Typically,

only the first 12–13 coefficients are retained, as they are considered sufficient to describe the overall shape of the spectral envelope, which is the key information required to model the vocal tract.

6. **Coefficient Selection**: After applying DCT, the first 2–13 coefficients are kept, while the higher-order coefficients are discarded as they contain less relevant information. These retained coefficients form the final MFCC features.

MFCC is employed as a primary feature to analyze cough sounds for the detection of COVID-19. Given its ability to capture the phonetic and spectral characteristics of sound, MFCC is well-suited for detecting subtle differences in cough patterns that may indicate COVID-19 infection.

3 Methods

This section details the methodology employed in the proposed COVID-19 cough detection system. We will explain the data preprocessing steps and the chosen models.

3.1 Data Preprocessing

The dataset we chose is COUGHVID [18] along with the COVID-19 Classification Dataset from Kaggle, containing 27,552 and 171 recording samples, respectively. The COUGHVID dataset includes 12,479 recording samples labeled as healthy, 2590 as symptomatic, 1155 as COVID-19, and 11,326 unlabeled recordings (see Table 1).

Table 1. Details of COUGHVID dataset.

Labels	Frequencies
Symptomatic	2,590
COVID-19	1,155
Healthy	12,479
Unlabeled	11,326

Apart from COUGHVID, for comparison purposes, we also use the COVID-19 Cough Recordings dataset from Kaggle which is still quite limited, comprising only 170 recordings, of which 151 are identified as non-COVID-19 cases and 19 as COVID-19 positive cases. The selection of this dataset is intended to provide a relevant basis for comparison with the COUGHVID dataset since these contain less complex datasets (see Table 2).

During our exploratory data analysis (EDA) of the COUGHVID dataset, we found that around eleven thousand entries had missing values for patient status. After careful consideration, we decided to exclude these rows from the dataset since imputing these missing values could introduce inaccuracies and potentially bias our results.

Table 2. Details of COVID-19 cough recordings dataset.

Labels	Frequencies
Non-COVID	151
COVID	19

Additionally, we have chosen to treat all Symptomatic samples in the COUGHVID dataset as COVID-19 samples. This decision simplifies model training and allows for easier comparisons between different datasets. To maintain consistency between datasets, we converted all audio to mono before extracting features. Specifically, we extracted the following features:

1. Chroma STFT (Short-Time Fourier Transform).
2. RMSE (Root Mean Square Energy).
3. Spectral Centroid.
4. Spectral Bandwidth.
5. Rolloff.
6. Zero-Crossing Rate (ZCR).

In addition, we include one of the most significant features in audio processing for machine learning is the Mel-Frequency Cepstral Coefficient (MFCC). MFCC is frequently employed in applications such as voice recognition, audio classification, and more recently in the medical field to extract essential audio features representing the vocal tract's shape. The process of computing MFCCs involves several steps, including framing the audio signal into short, overlapping frames, calculating the power spectrum for each frame, applying the Mel Filter Bank to the power spectra, taking the logarithm of the filter bank energies, and performing the Inverse Discrete Cosine Transform (IDCT) on the log filter bank (see Fig. 1) [22].

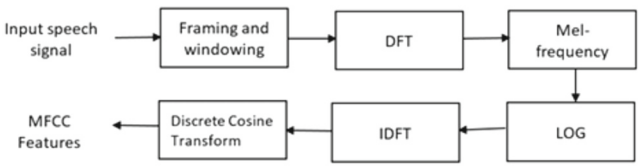


Fig. 1. MFCC process diagram

After the extraction of audio features, we perform feature scaling to convert the 2D array of MFCC features into a 1D array. This transformation involves calculating the mean MFCC coefficient for each respective frame over the entire audio file, thus reducing the dimensionality while preserving crucial information.

To prepare the extracted data for training, we split it into two sets: a training set and a test set with an 80:20 ratio. We then applied Mahalanobis Distance-based Oversampling (MDO) to the minority class (COVID-19) in the training set to address the imbalance between label classes. The Mahalanobis Distance-based Over-sampling Technique

(MDO) generates new synthetic samples that maintain the same Mahalanobis distance from the class mean as the existing minority class examples [23, 24]. This technique helps to mitigate the dominance of the majority class (healthy) and ensures a more balanced representation of both classes during training.

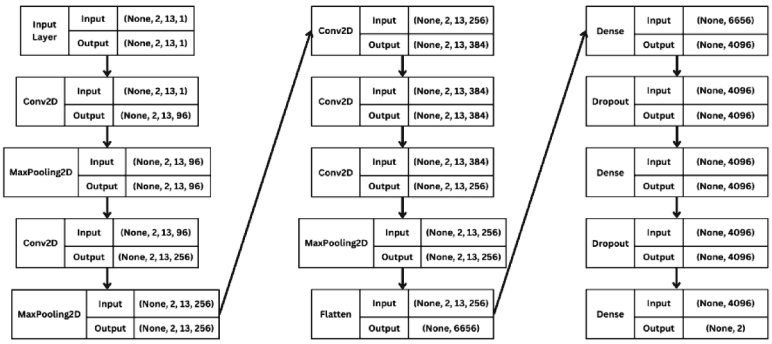
3.2 Models

XGBoost. eXtreme Gradient Boosting (XGBoost) is an advanced implementation of gradient boosting that incorporates several enhancements, including, Decision Trees as Base Learners, Objective Function, Regularization, Shrinkage (Learning Rate), Parallel and Distributed Computing, Sparsity-aware, Cross-validation, and Early Stopping. These enhancements make XGBoost one of the best methods for classifying tabular datasets, often outperforming more complex methods such as neural networks.

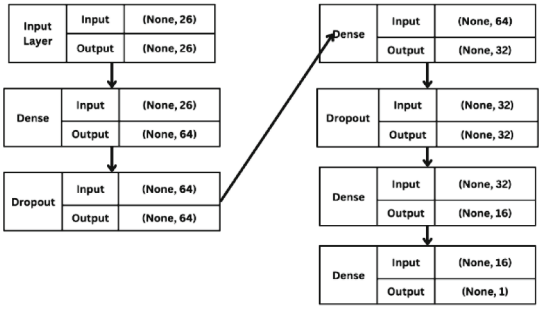
Random Forests. Random Forests are a powerful and widely used ensemble learning method that combines multiple decision trees to improve classification accuracy and robustness [25]. The main advantage of Random Forests lies in their ability to enhance predictive accuracy and reveal feature importance. By aggregating the votes from individual trees, Random Forests provides a final class prediction that is typically more accurate than that of any single tree. This method not only delivers high predictive performance but also allows for the identification of crucial features and their interactions.

AlexNet. AlexNet is a Convolutional Neural Network (CNN) architecture developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton. It was designed for image classification tasks and gained prominence after winning the ImageNet Large Scale Visual Recognition Challenge in 2012. The network is composed of: Five Convolutional Layers, Three Max Pooling Layers, Two Normalization Layers, Three Fully Connected Layers, Dropout Layers, ReLU Activation Functions. AlexNet introduced several innovations such as the ReLU activation function and dropout layer. These advancements proved to be crucial to the model training process, making AlexNet a foundation for many modern neural networks used today [26]. Our proposed AlexNet architecture can be seen in Fig. 2a.

FCNN. Fully connected deep networks, also known as fully connected neural networks (see Fig. 3a), are fundamental tools in deep learning, applicable to a wide variety of problems. Their primary advantage is that they do not require any assumptions about the input data's structure, making them versatile for different data types. These networks consist of multiple layers, often called fully connected layers, where each layer's neurons are connected to every neuron in the previous layer, forming a complex network capable of learning intricate patterns. The fully connected neural network model (see Fig. 3b) we have opted to use is designed for binary classification tasks. It includes an input layer that processes input vectors with 26 features, followed by three hidden layers with ReLU activations and dropout layers for regularization. The final output layer uses a sigmoid activation function to produce a binary outcome. This architecture leverages dropouts to prevent overfitting, ensuring that the model generalizes well to new data. The model is trained using the Adam optimizer, which efficiently handles the optimization process by adapting the learning rate. Additionally, early stopping is implemented to halt training once the model's performance on a validation set stops improving, further safeguarding against overfitting [27]. Our proposed FCNN architecture can be seen in Fig. 2b.



(a)



(b)

Fig. 2. Architecture of (**a**) AlexNet, (**b**) FCNN

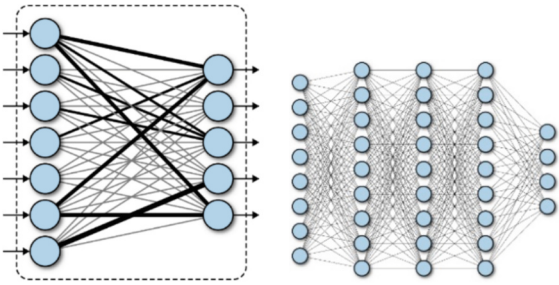


Fig. 3. (**a**) Fully connected layer diagram, (**b**) Fully connected neural network diagram

3.3 Model Training

The model was trained using preprocessed datasets. In ensemble learning, each method's hyperparameters were tested to achieve optimal results. We conducted experiments using 3-fold cross-validation to prevent model overfitting. All those parameters were tuned using Grid Search Cross Validation with 3-fold cross-validation. Grid Search is a method

that tests various combinations of parameters to find the best model. It creates and evaluates a model for every possible combination and selects the most accurate one. However, this technique can be time-consuming, as it considers all possible combinations, so the run time depends on the number of parameters and their possible values [28]. We use Grid Search CV due to its numerous implementations [29] and its proven effectiveness with XGBoost and Random Forest, as demonstrated in previous studies [30, 31]. For XGBoost, we tuned the following parameters with the following choices (see Table 3).

Table 3. Details of XGBoost hyperparameter tuning.

Parameters	Parameter choices	Best parameter (COUGHVID)	Best parameter (COVID-19 cough recordings)
n_estimators	100, 200, 300	300	300
max_depth	3, 5	5	5
learning_rate	0.01, 0.001	0.01	0.01
subsample	0.6, 0.8, 1.0	0.6	0.6
colsample_bytree	0.6, 0.8, 1.0	1.0	0.6
gamma	0, 0.1, 0.2, 0.3	0.2	0
reg_alpha	0, 0.1, 0.5, 1.0	0.1	0.5
reg_lambda	0, 0.1, 0.5, 1.0	0.1	0.1

For the COUGHVID Dataset, the optimal parameters determined are as follows: 'n_estimators': 300, 'max_depth': 5, 'learning_rate': 0.01, 'subsample': 0.6, 'colsample_bytree': 1.0, 'gamma': 0.2, 'reg_alpha': 0.1, and 'reg_lambda': 0.1. Meanwhile, for the COVID-19 Cough Recordings Dataset, we get 'n_estimators': 300, 'max_depth': 5, 'learning_rate': 0.01, 'subsample': 0.6, 'colsample_bytree': 0.6, 'gamma': 0, 'reg_alpha': 0.5, 'reg_lambda': 0.1. Whilst in Random Forest, the parameters and their choices are as follows (see Table 4).

Table 4. Details of random forest hyperparameter tuning.

Parameters	Parameter choices	Best parameter (COUGHVID)	Best parameter (COVID-19 cough recordings)
n_estimators	25, 50, 100, 150	50	150
max_features	sqrt, log2, None	log2	Log2
max_depth	3, 6, 9	6	6
max_leaf_nodes	3, 6, 9	9	9

For the COUGHVID Dataset, the best parameters identified are: 'n_estimators': 50, 'max_features': 'log2', 'max_depth': 6, and 'max_leaf_nodes': 9. In contrast, for the COVID-19 Cough Recordings Dataset, the optimal parameters are: 'n_estimators': 150, 'max_features': 'log2', 'max_depth': 6, and 'max_leaf_nodes': 9.

In the Neural Network, during the model training process, we implement early stopping to prevent overfitting and reduce the risk of excessive training, which can cause

the model to overly fit the training data and lose its ability to generalize well on new data [32]. We conducted training for 100 epochs with a batch size of 32, as this batch size allows for more stable and efficient training by updating the model weights more frequently compared to larger batches. This approach helps to strike a balance between training speed and result quality, enabling the model to learn from the data in a controlled and optimal manner.

4 Results

In this section, we present the results of our experiments and provide a detailed discussion of the findings. Our goal is to evaluate the performance of various deep learning models on the COUGHVID dataset for the classification of COVID-19 and healthy cough sounds.

After completing the modeling process, each model will be thoroughly evaluated to ensure their effectiveness and reliability. Given the imbalance in the dataset, it is essential to consider both accuracy and the F1 score for each label to obtain a comprehensive understanding of the models' performance. Evaluating the models on these metrics will help assess their overall correctness and ability to identify positive and negative cases, which is crucial in handling imbalanced data. Below, we present the evaluation table for the four models across both datasets, providing a detailed comparison of their performance metrics (see Table 5).

Table 5. Results.

Models	COUGHVID dataset			COVID-19 cough recordings dataset		
	F1-score on healthy (%)	F1-score on COVID-19 (%)	Accuracy (%)	F1-score on healthy (%)	F1-score on COVID-19 (%)	Accuracy (%)
XGBoost	85	10	75	96	77	93
Random Forest	83	16	72	93	62	88
FCNN	86	3	76	97	83	95
AlexNet	86	2	76	96	77	93

To further enhance our understanding, we also analyze the results using a confusion matrix. This approach provides a clear visualization of the model's performance by showing the true positives, true negatives, false positives, and false negatives, enabling us to better interpret the effectiveness of our models.

In Fig. 4, the confusion matrix shows that models perform better on the Kaggle dataset than on the COUGHVID dataset, likely due to imbalanced data and differences in the datasets. XGBoost and Random Forest perform well on Kaggle but poorly on COUGHVID, while FCNN and AlexNet also do well on Kaggle but struggle with COUGHVID.

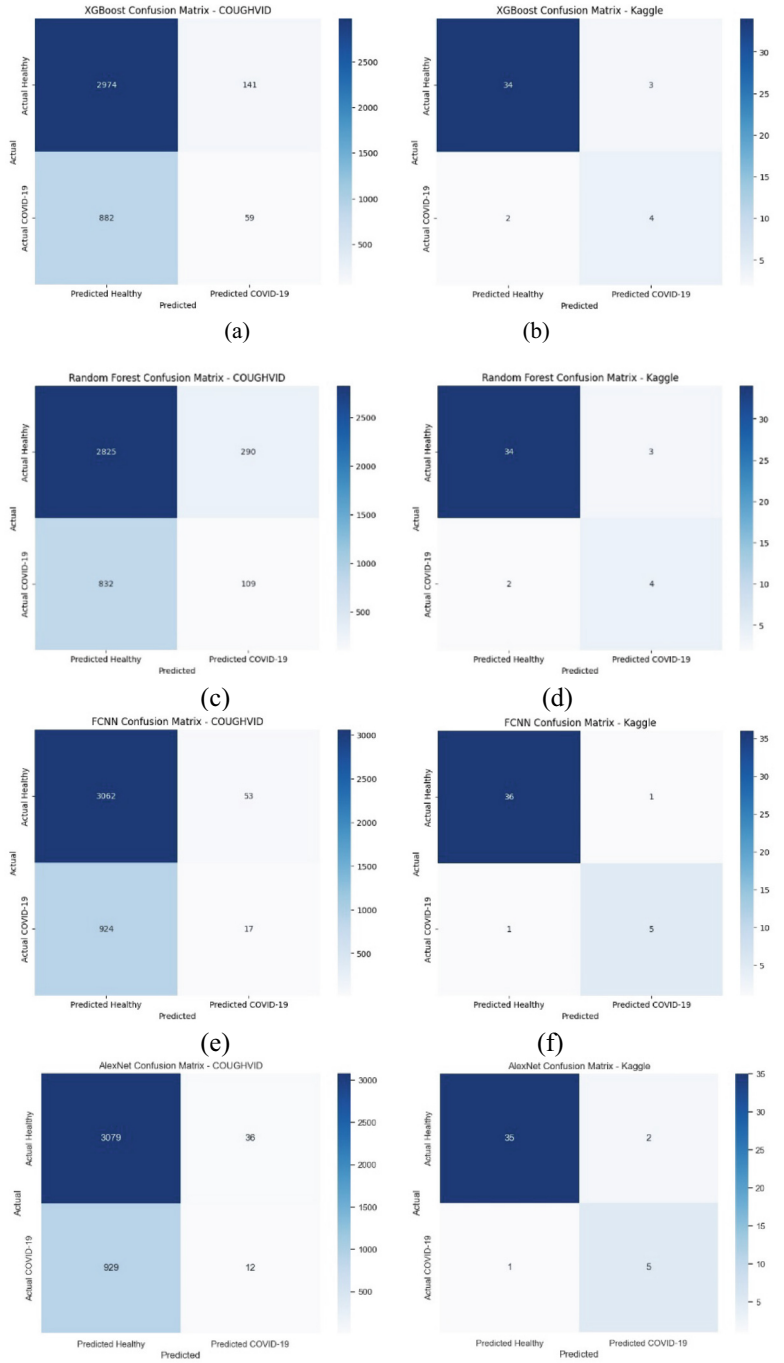


Fig. 4. Confusion Matrix of (**a**) XGBoost on COUGHVID Dataset, (**b**) XGBoost on Kaggle Dataset, (**c**) Random Forest on COUGHVID Dataset, (**d**) Random Forest on Kaggle Dataset, (**e**) FCNN on COUGHVID Dataset, (**f**) FCNN on Kaggle Dataset, (**g**) AlexNet on COUGHVID Dataset, (**h**) AlexNet on Kaggle Dataset

5 Discussions

Model Evaluation Analysis on the COUGHVID Dataset. The evaluation results indicate that all models perform ineffectively on the COUGHVID Dataset, as the highest F1-scores for all models towards COVID-19 label was attained by Random Forests at 16%, which is still below what we consider acceptable. Although the accuracy on the COUGHVID Dataset tends to be high, this is due to the considerable number of healthy data, making the predictions for healthy conditions more likely to be correct. The accuracy of each model tends to be similar, averaging around 75%. These results suggest that all models do not work effectively on the imbalanced dataset with a large amount of data.

Model Evaluation Analysis on the COVID-19 Recordings Dataset. Model evaluation shows that the FCNN model performed the best with an accuracy of 95%, with AlexNet and XGBoost trailing only shortly behind at 93%. Although evaluation results show higher accuracy and F1-scores compared to the COUGHVID Dataset, it is important to note that this dataset only consists of 170 data points and is not exceptionally large. Each model has an average accuracy of 92%, which is high, but we should not assume that these models are robust given the dataset's limitations.

Overall Analysis. The evaluation results indicate that the F1-scores for each label are similar across different datasets, despite using different models. Therefore, we think that preprocessing and handling of imbalanced data are also one of the factors contributing to the F1-spercore and accuracy, as they significantly impact the final F1-score and accuracy. So far, none of the four models can be considered state-of-the-art on imbalanced datasets, even after being handled using MDO Resampling.

The results indicate significant variability in model performance across different datasets, highlighting the challenges of working with imbalanced datasets. The Random Forest model shows relatively better performance on the COUGHVID dataset in terms of F1-score for COVID-19 detection, which is 16%, compared to other models, yet the score is still below acceptable standards for clinical applications. This suggests a limitation in the current machine learning models' ability to handle datasets where one class overwhelmingly outnumbers another.

On the COVID-19 Cough Recordings Dataset, the FCNN model shows the best performance, achieving an accuracy of 95% and F1-scores of 97% for the healthy label and 83% for the COVID-19 label. This suggests that FCNN might be better suited for smaller imbalanced datasets compared to other models like XGBoost, Random Forest, or AlexNet. However, given the small size of this dataset, these results might not generalize well to larger, real-world datasets.

In comparison to other studies, using MFCC features achieved an accuracy of 70.58%, which is close to our results, 72–76% accuracy in large datasets. However, since their approach did not use actual data, our use of actual data has slightly increased the accuracy [33]. As for model performance comparison, it can be seen through Table 6, and since in the real world, the large dataset is more relevant, we will use our accuracy in COUGHVID Dataset.

Table 6. Comparison with other related studies.

Model	Accuracy (%)
XGBoost (ours)	75
Random Forest (ours)	72
FCNN (ours)	76
AlexNet (ours)	76
Random Forest [1]	91.92
AlexNet [12]	86
BiGRU-Attention-XGBoost [19]	72.45
Deep Neural Network (DNN) [9]	96
HMM-GMM-MFCC [11]	91.57
VGG-19 [13]	89
CNN-LSTM [34]	91.13
C3-PO [35]	72.3
SL-3 [36]	78.85

Table 6 illustrates the accuracy of various machine learning models, including those proposed by this paper (indicated as "ours") and those sourced from the literature. Notably, our Random Forest model exhibits the lowest accuracy at 72%, whereas the Deep Neural Network (DNN) in [9] achieves the highest accuracy at 96%. It's clear that our methods may not be the state-of-the-art method in the field, as several models demonstrate accuracies above 90%. Additionally, the recently proposed models, C3-PO and SL3 from early 2024, highlight that newer models may require time to optimize and improve their accuracy levels. This perspective is crucial for understanding the developmental trajectory and potential of emerging machine learning techniques.

The limitations of our study are underscored by several factors, including the size, which is only around 170 samples in the small dataset) and composition of the dataset used, in which the large dataset contains healthy samples that are more than three times the combined total of symptomatic and COVID-19 samples. The COVID-19 Cough Recordings Dataset's small size curtails the generalizability of the findings, as larger datasets could provide different insights and expose the models to greater variability. Furthermore, the imbalance in data presents another challenge, as the models show high performance in one class but not in another. This suggests that these models may not be robust enough for clinical applications that demand high accuracy across all classes.

The results of our study are pivotal for advancing machine learning applications in healthcare, particularly in developing diagnostic tools for infectious diseases like COVID-19. Our findings highlight both the potential and challenges associated with current machine learning models when applied to medical datasets, especially those with significant class imbalances. In healthcare, where precision is crucial, our study emphasizes the need to address the issues posed by imbalanced datasets. These imbalances are common in real-world medical scenarios, where certain conditions, such as healthy states, far outnumber others, such as specific diseases.

Our study underscores the importance of refined preprocessing techniques to manage imbalanced datasets more effectively. Robust feature selection methods are also crucial for enhancing model accuracy and reliability. Metrics beyond mere accuracy, such as the F1-score for each class, offer a more comprehensive view of a model's performance. Relying solely on accuracy can be misleading, especially in datasets dominated by one class, potentially obscuring the model's ability to detect critical cases like COVID-19 infections.

Advanced preprocessing techniques, including sophisticated resampling methods and data augmentation strategies, are essential for mitigating the effects of imbalance. Additionally, feature selection methods that prioritize the most informative aspects of the data can significantly improve model robustness and generalizability, making them more effective in diverse clinical settings.

As healthcare increasingly relies on data-driven decisions, it is crucial that machine learning models are developed and evaluated to ensure their fairness and effectiveness across various diseases and patient demographics. Testing these models on datasets that reflect the complexity of real-world scenarios is key to their success. Furthermore, ongoing evaluation and refinement of these models are necessary to adapt to new data and evolving medical knowledge, ensuring their continued relevance and efficacy as diagnostic tools.

6 Conclusion

The aim of this study is to provide insights that can guide researchers and practitioners in selecting the most suitable machine learning model for their specific needs, contributing to more effective and efficient applications in various fields.

Our findings highlight the critical role of data preprocessing in model performance. For instance, selecting appropriate features from audio data is essential and requires careful experimentation. In terms of performance metrics, the F1-score was used to provide a clearer picture of model efficacy on imbalanced datasets. Observing through F1-score for the COVID label, in the case of large imbalanced datasets, ensemble learning models, especially Random Forest, achieved an F1-score of 16%, which is higher than that of neural networks, which had a maximum F1-score of 3% in FCNN. Conversely, for smaller imbalanced datasets, neural networks outperformed ensemble learning models, with FCNN reaching an F1-score of 83%, whereas ensemble methods like XGBoost had a highest F1-score of 77%.

Future research should focus on identifying optimal features and evaluation techniques and exploring additional models beyond the two neural networks and two ensemble models assessed in this study. Additionally, efforts should be directed toward acquiring more balanced and higher-quality datasets to enhance model performance and reliability. By addressing these challenges, researchers and practitioners can improve model selection and application, ultimately enhancing the effectiveness and efficiency of machine learning solutions in various domains.

Acknowledgement. We would like to extend our gratitude to the Embedded Systems Laboratory (ESL) of the Institute of Electrical and Micro Engineering at EPFL for providing the COUGHVID dataset. We also thank Mr. Himanshu for the COVID-19 Cough Recordings Dataset from Kaggle.

References

1. Tena A, Clarià F, Solsona F (2022) Automated detection of COVID-19 cough. Biomed Signal Process Control 71:103175. https://doi.org/10.1016/J.BSPC.2021.103175
2. Gong W, Parkkila S, Wu X, Aspatwar A (2023) SARS-CoV-2 variants and COVID-19 vaccines: current challenges and future strategies. Int Rev Immunol 42:393–414. https://doi.org/10.1080/08830185.2022.2079642
3. Ulukaya S, Sarıca AA, Erdem O, Karaali A (2023) MSCCov19Net: multi-branch deep learning model for COVID-19 detection from cough sounds. Med Biol Eng Comput 61:1619–1629. https://doi.org/10.1007/S11517-023-02803-4/TABLES/7
4. COVID-19 cases|WHO COVID-19 dashboard. https://data.who.int/dashboards/covid19/cases?n=o. Accessed 7 Oct 2024
5. Artika IM, Dewi YP, Nainggolan IM, Siregar JE, Antonjaya U (2022) Real-time polymerase chain reaction: current techniques, applications, and role in COVID-19 diagnosis. Genes (Basel) 13:2387. https://doi.org/10.3390/GENES13122387
6. Alghamdi HS, Amoudi G, Elhag S, Saeedi K, Nasser J (2021) Deep learning approaches for detecting COVID-19 from chest x-ray images: a survey. IEEE Access 9:20235–20254. https://doi.org/10.1109/ACCESS.2021.3054484
7. Fu F et al (2020) Chest computed tomography findings of coronavirus disease 2019 (COVID-19) pneumonia. Eur Radiol 30:5489–5498. https://doi.org/10.1007/S00330-020-06920-8
8. Quer G et al (2021) Wearable sensor data and self-reported symptoms for COVID-19 detection. Nat Med 27:73–77. https://doi.org/10.1038/S41591-020-1123-X
9. Kim S, Baek JY, Lee SP (2023) COVID-19 detection model with acoustic features from cough sound and its application. Appl Sci 13:2378. https://doi.org/10.3390/APP13042378
10. Mafazy MM, Faisal MR, Kartini D, Indriani F, Saragih TH (2023) Classification of COVID-19 cough sounds using Mel frequency cepstral coefficient (MFCC) feature extraction and support vector machine. Telematika 16:60–68. https://doi.org/10.35671/TELEMATIKA.V16I2.2569
11. Hamidi M, Zealouk O, Satori H, Laaidi N, Salek A (2023) COVID-19 assessment using HMM cough recognition system. Int J Inf Technol 15:193–201. https://doi.org/10.1007/S41870-022-01120-7/FIGURES/14
12. Nafiz MF, Kartini D, Faisal MR, Indriani F, Saragih TH (2023) Automated detection of COVID-19 cough sound using mel-spectrogram images and convolutional neural network. J Ilmiah Teknik Elektro Komput Inf 9:535–548. https://doi.org/10.26555/JITEKI.V9I3.26374
13. Effati M, Nejat G (2023) A performance study of CNN architectures for the autonomous detection of COVID-19 symptoms using cough and breathing. Computers 12:44. https://doi.org/10.3390/COMPUTERS12020044
14. Södergren I, Nodeh MP, Chhipa PC, Nikolaidou K, Kovács G (2021) Detecting COVID-19 from audio recording of coughs using Random Forests and Support Vector Machines. In: Proceedings of the annual conference of the international speech communication association, Interspeech, vol 6, pp 916–920. https://doi.org/10.21437/INTERSPEECH.2021-2191
15. Rahman DA, Lestari DP (2021) COVID-19 classification using cough sounds. In: Proceedings—2021 8th international conference on advanced informatics: concepts, theory, and application, ICAICTA 2021. https://doi.org/10.1109/ICAICTA53211.2021.9640278
16. Sevi M, Aydin I (2020) COVID-19 detection using deep learning methods. In: 2020 International conference on data analytics for business and industry: way towards a sustainable economy, ICDABI 2020. https://doi.org/10.1109/ICDABI51230.2020.9325626
17. Bhoumik S, Chatterjee S, Sarkar A, Kumar A, Joseph FJJ (2020) Covid 19 Prediction from X Ray images using fully connected convolutional neural network. In: ACM International conference proceeding series, pp 106–107. https://doi.org/10.1145/3429210.3429233

18. Gupta P, Degadwala S (2023) A comprehensive review on COVID-19 cough audio classification through deep learning. Int J Sci Res Comput Sci Eng Inf Technol. https://doi.org/10.32628/CSEIT2361049
19. Zhao X, Shao Y, Mai J, Yin A, Xu S (2020) Respiratory sound classification based on BiGRU-attention network with XGBoost. In: Proceedings—2020 IEEE international conference on bioinformatics and biomedicine, BIBM 2020, pp 915–920. https://doi.org/10.1109/BIBM49941.2020.9313506
20. Dar JA, Srivastava KK, Ahmed Lone S (2022) Design and development of hybrid optimization enabled deep learning model for COVID-19 detection with comparative analysis with DCNN, BIAT-GRU. XGBoost Comput Biol Med 150:106123. https://doi.org/10.1016/J.COMPBIOMED.2022.106123
21. Grant D, McLane I, West J (2021) Rapid and scalable COVID-19 screening using speech, breath, and cough recordings. In: BHI 2021—2021 IEEE EMBS international conference on biomedical and health informatics, proceedings. https://doi.org/10.1109/BHI50953.2021.9508482
22. Abdul ZK (2019) Kurdish speaker identification based on one dimensional convolutional neural network. In: Computational methods for differential equations, vol 7, pp 566–572
23. Altwaijry N (2023) Probability-based synthetic minority oversampling technique. IEEE Access 11:28831–28839. https://doi.org/10.1109/ACCESS.2023.3260723
24. Pristyanto Y, Mukarabiman Z, Nugraha AF (2023) Extreme gradient boosting algorithm to improve machine learning model performance on multiclass imbalanced dataset. Int J Inf Vis 7:710–715. https://doi.org/10.30630/JOIV.7.3.1102
25. Denisko D, Hoffman MM (2018) Classification and interaction in random forests. Proc Natl Acad Sci USA 115:1690. https://doi.org/10.1073/PNAS.1800256115
26. Lu S, Lu Z, Zhang YD (2019) Pathological brain detection based on AlexNet and transfer learning. J Comput Sci 30:41–47. https://doi.org/10.1016/J.JOCS.2018.11.008
27. Zadeh RB, Ramsundar B (2018) TensorFlow for deep learning from linear regression to reinforcement learning
28. Hansen CB, Mahmoud Ali AM, Ethelberg MZ, Wozniak MM (2019) Comparing Random Forest, XGBoost and neural networks with hyperparameter optimization by nested cross-validation
29. Shabrandi AR, Rajabzadeh Ghatari A, Tavakoli N, Dehghan Nayeri M, Mirzaei S (2023) Fast COVID-19 infection prediction with in-house data using machine learning classification algorithms: a case study of Iran. J AI Data Mining 11:573–585. https://doi.org/10.22044/JADM.2023.13291.2458
30. Hamida S, Gannour OEL, Cherradi B, Ouajji H, Raihani A (2020) Optimization of machine learning algorithms hyper-parameters for improving the prediction of patients infected with COVID-19. In: 2020 IEEE 2nd international conference on electronics, control, optimization and computer science, ICECOCS 2020. https://doi.org/10.1109/ICECOCS50124.2020.9314373
31. Baita A, Prasetyo A, Cahyono N, Yogyakarta A (2023) Hyperparameter tuning on random forest for diagnose COVID-19. JIKO. https://doi.org/10.33387/JIKO.V6I2.6389
32. Keles A, Keles MB, Keles A (2024) COV19-CNNet and COV19-ResNet: diagnostic inference engines for early detection of COVID-19. Cogn Comput 16:1612–1622. https://doi.org/10.1007/S12559-020-09795-5/FIGURES/6
33. Bansal V, Pahwa G, Kannan N (2020) Cough classification for COVID-19 based on audio mfcc features using convolutional neural networks. In: 2020 IEEE international conference on computing, power and communication technologies, GUCON 2020, pp 604–608. https://doi.org/10.1109/GUCON48875.2020.9231094

34. Hamdi S, Oussalah M, Moussaoui A, Saidi M (2022) Attention-based hybrid CNN-LSTM and spectral data augmentation for COVID-19 diagnosis from cough sound. J Intell Inf Syst 59:367–389. https://doi.org/10.1007/S10844-022-00707-7
35. Chen X, Khan MAR, Hasan MR, Gedeon T, Hossain MZ (2024) C3-PO: a convolutional neural network for COVID onset prediction from cough sounds. In: Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and lecture notes in bioinformatics). 14556 LNCS, pp 355–368. https://doi.org/10.1007/978-3-031-53311-2_26
36. Topuz EK, Kaya Y (2024) SUPER-COUGH: a super learner-based ensemble machine learning method for detecting disease on cough acoustic signals. Biomed Signal Process Control 93:106165. https://doi.org/10.1016/J.BSPC.2024.106165

Model Development and Features Selection of Brute Force Attack Classification on Application Layer Network

Muhammad Afrizal Amrustian[1(✉)], Bita Parga Zen[2], and Muhammad Ikhwani Saputra[3]

[1] Institut Teknologi Telkom Purwokerto, Purwokerto, Indonesia
afrizal.amru@ittelkom-pwt.ac.id
[2] Universitas Ma Chung, Malang, Indonesia
[3] Universitas Siber Asia, Jakarta, Indonesia

Abstract. A brute force attack is a kind of attack on the internet that aims for personal information. Even though brute force is an old attack, but it can crack the encrypted data. The research has been conducted on identifying Brute Force attack, but they did not focus on the application layer of network. Therefore, this research aims on identifying brute force attacks on application layer. The CICIDS 2017 dataset was used and consists of three kinds of label network traffic, normal traffic, FTP Patator, and SSH Patator as brute force attack traffic. Decision trees and random forest are the classifiers used in this research. The research results are both of decision trees and random forests reach 99.9% accuracies. The precision result for the decision tree is 99.9% for normal traffic, 99.9 for FTP Patator, and 99.8 for SSH Patator. The recall result for the decision tree is 99.9% for normal, 99.9% for FTP Patator, and 100% for SSH Patator. The precision for the random forest is 100% for normal, 100% for FTP brute Patator, and 99.7% for SSH Patator. The recall for the random forest is 99.9% for normal, 100% for FTP Patator, and 100% for SSH Patator. Although the performance did not differ significantly, the decision tree has the advantage over the random forest in running time. The decision tree is seven seconds faster than the random forest. Furthermore, this research contributes in defining five importance features to recognize the brute force attack. The five features are Destination Port, Maximal Forward Packet Length, Backward Init Win Bytes, Mean Packet Length, and Average Packet Size.

Keywords: Brute Force · Classification · Decision Tree · Random Forest

1 Introduction

People can easily do anything using the internet in the digital era, such as searching for information, transferring money, and shopping. However, those activities involve personal information, and people often lack awareness about keeping their personal information [1–3]. One of the attacks that are looking for personal information is a brute force attack [4–6]. According to the 2021 Data Breach Investigations Report and news

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 109–120, 2026.
https://doi.org/10.1007/978-981-95-9233-3_9

from the mission-critical website, brute force is still a major attack on the internet [7, 8]. The brute force attack also costs Tesco, one of the biggest supermarkets in the UK, by breaking the server and making more than a thousand new club cards [9].

Brute force is an attack that cracks the password by trying all possible combinations [10–13]. A brute force attack is an old attack method, but the attack still can break the security by trying all possible usernames and password [14], and brute force even can crack the encrypted password or ssh login. Brute force attacks can be more sophisticated and faster in the cracking system by using a high-performance computer. Ju Jo and Won Yoo have proved that brute force attack with a high-performance computer can break the system faster [15]. Therefore, brute force attack is considered as one of the distressing attacks. Some researchers have been done overcome brute force attacks. Amijoyo et al. classified brute force attacks using the Naïve Bayes method. That research used Hydra as a tool to launch brute force attacks and Wireshark to capture the traffic. The result is naïve Bayes can classify the brute force attack, but the accuracy only reaches 50% [16]. Mesran and Nasution combined caesar cipher and stout codes to encrypt the data [17]. The result is that brute force cannot break the encryption. However, combining two cryptography methods is not commonly use, and hard to implement the method. The use of multiple encryption techniques was proposed by Kushwaha et al., and the idea is to use more than one encryption to prevent brute force attacks [18]. The result is not clear enough to prevent from brute force attack. Wanjau et al. are used a convolutional neural network (CNN) to determine SSH brute force attacks. CNN able to detect SSH brute force attack with 94.4% accuracy, 92.5% precision, 97.8% recall and 91.8 F1-score [19]. Moreover, the running time is not declared in this research. Sandra et al. visualized the brute force attack using k-means and Naïve Bayes. Based on the visualization, Sandra et al. can determine the pattern from brute force attacks. K-means reaches 95.46% accuracy, and Naïve Bayes reaches 73.60% [20].

Nathiya and Suseendran compared three methods Support Vector Machine (SVM), Naïve Bayes, and Decision Tree for detecting malicious traffic, including brute force attack. This research uses the NSL-KDD dataset, and the result is that the Decision Tree gains the highest accuracy compared to SVM and Naïve Bayes [21]. Negandhi et al. used the random forest as a classifier to detect several attacks in a traffic network. The result has shown that random forest can classify various kinds of attacks in the NSL-KDD dataset and reach 99.7% accuracy [22]. According to several existing researchers, Naive Bayes can not be used as a method for classifying brute force attacks because of the lack of accuracy. Deep learning methods like CNN achieve high accuracy in classifying brute force attacks, but there is no information about the running time when CNN is used to detect brute force attacks. Data encryption is considered as an option to overcome brute force attacks, but the method is not quite popular to use. For example, the decision tree and random forest can classify attacks in the NSL KDD dataset correctly. However, the used dataset is the old dataset and not focused on brute force attacks. Therefore, this research focuses on brute force attack detection. From the conclusion, this research compares decision trees and random forests to classify brute force attacks based on the problem and existing research. Using the CICID IDS 2017 dataset and focusing on the

brute force attack dataset. This paper consists of the introduction, the research method, the result and discussion, and the conclusion. Section 2 is explaining about research method.

2 Methods

The research method consists of five steps. The first step is collecting the data. Then, the second is data cleaning, and the third is data splitting. The fourth step is classification, and the fifth step is evaluation. The description of the research method can be seen in Fig. 1.

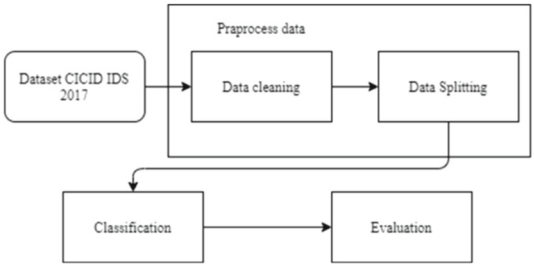


Fig. 1. Research method

2.1 Dataset

The dataset used in this paper was taken from the University of New Brunswick (UNB). UNB has various kinds of network attack datasets. UNB creates all network attack datasets by simulating all types of attacks and capturing them. The network attack dataset provided by UNB was divided into two types. The first is .pcap as the raw file. The second is the .csv type. This research uses the .csv file type to simplify preprocessing data. The name of the dataset is CICIDS 2017. The dataset consists of 78 network features and labels with three classes benign class as normal network, SSH patator, and FTP patator as attack network [23]. The benign class consists of some network protocols like HTTP, HTTPS, FTP, SSH, and email protocols. UNB uses the patator to launch brute force attacks and attack two protocols, FTP and SSH, for the attack class.

2.2 Data Cleaning

Before classifying data, the data must through the preprocessing step, and data cleaning is the first step in preprocessing data. The data cleaning is the step to search for some missing value or other data that will lead to errors in classification as shown below on Fig. 2 [24, 25]. One technique for cleaning the data is litwise deletion. This approach deletes the dataset's missing value, but the data remains must be large after wiping

the data [26]. This research uses litwise deletion because the remaining data still large after deletion of some missing value. The algorithm below is shown the process of data cleaning.

The data cleaning process will eliminate two types of data: NaN data and the second is unlimited data. Unfortunately, Na and infinite data can not be calculated. Hence, deletion is a better option.

```
Data cleaning algorithm
Input: Network data.csv
Output: Cleaned network data.csv
  Initialization data
  Read data <- network data.csv sep = ,
  data <- na.omit
  data <- inf.omit
```

Fig. 2. Algorithm of data cleaning

2.3 Data Splitting

Data splitting is the step to divide the dataset into two parts, the first is the training part, and the second is the test part. The training part is used for building the classification model and the test part for testing the model's performance. Thus, the dataset will split into 70% for training data and 30% for testing data. The algorithm in Fig. 3 is shown the process of splitting data into training and testing data.

```
Data splitting algorithm
Input: data
Output: training data, testing data
  Initialization sample
  Seet seed
  Sample      <-      sample.split(data$label,
splitratio = .70)
  Training  data  <-  subset(data,  sample  ==
TRUE)
  Testing  data  <-  subset(data,  sample  ==
FALSE)
```

Fig. 3. Algorithm of data splitting

2.4 Classification

Decision Tree

The decision tree is a classification method that changes the data into some rules of decision. The decision tree consists of a root node, internal node, and leaf node [27–30]. Figure 4 is the structure of the decision tree. Decision tree consist of Entropy and Information Gain calculation. The entropy is the mathematical formula for measuring

the dataset randomness. The information gain has the function to informs how much knowledge of a random feature's value. The entropy is calculated by Eq. (1) and the information gain is obtained with Eq. (2) [30].

$$Entropy(S) = \sum_{i=1}^{c} P_i \log 2^{Pi} \tag{1}$$

$$Gain(S, A) = \sum_{v \in V(A)} \frac{|S_v|}{|S|} Entropy(S_v) \tag{2}$$

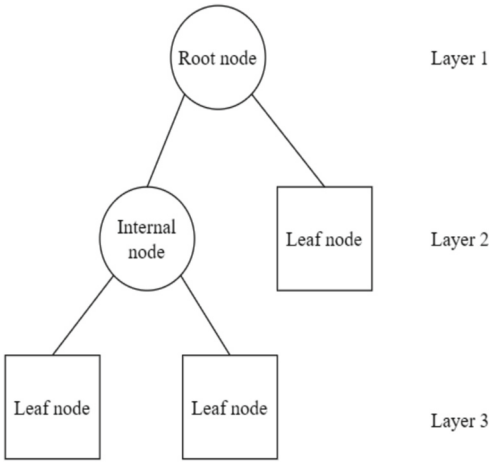


Fig. 4. Decision tree structure [31]

The root node is the first rule in the decision tree that will split the data, in this part using the training data that have 70% from all data, into the next step. The internal node is the middle node with input from the node which also works as the next rule. Therefore, the leaf node is the result of the decision tree. The algorithm below explains the building model process (Fig. 5). The training data will be used in the model development process.

```
Decision tree model process
Input: training data
Output: decision tree model classification
  require j48
   Decision_tree_model <- J48(class~., data =
training data)
```

Fig. 5. Process of decision tree model

The model classification was build from training data. J48 was chosen as a package to build a decision tree model. Based on class as a guide to determine the class for each data.

Random Forest

Random forest is an ensemble method for classification. The random forest consists of some decision trees. The decision tree will give an output for classification, and the last after all tree gives the output, the random forest will do the vote to take the last result of the classification [32–34]. Figure 6 is a depicted random forest classification architecture. In Fig. 6, tree1, tree2, …, treeB is the individual decision tree in the random forest method. k1, k2, …, kB are the output for each decision tree. Finally, K is the final result from random forest classification.

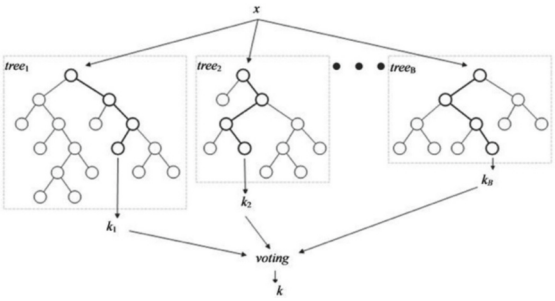


Fig. 6. The architecture of random forest

The algorithm for building a random forest model is shown in the algorithm below. The process uses training data as a guide to classify each class from each data (Fig. 7). The Random forest package was the package for building a random forest classification model.

```
Random forest model process
Input: training data
Output: random forest model classification
  require randomforest
  random_forest_model <- randomforest(class~.,
data = training data)
```

Fig. 7. Process of random forest model

Evaluation

Evaluation is the step to examine the performance of the classification model. This paper uses accuracy, precision, and recall to evaluate model performance. These three indicators are calculated from the confusion matrix, the confusion matrix shown in Table 1.

Table 1. Confusion matrix

		True class	
		Positive	Negative
Predicted class	Positive	True positive (TP)	False negative (FN)
	Negative	False positive (FP)	True negative (TN)

Where TP is the normal traffic is classified rightly. FP is the brute force traffic were classified as normal traffic. FN is the normal traffic that was classified as brute force traffic. TN is the brute force traffic were classified rightly. The equation for accuracy, precision, and recall calculation is shown in Eqs. (3), (4) and (5) [35]. The accuracy is the calculation to see how much the model can correctly classify brute force and normal traffic. The accuracy calculation is shown in Eq. (3). Precision is the proportion of the positive identifications that were correct, and the calculation is shown in Eq. (4). Besides, the recall is that the actual positives were identified correctly, and Eq. (5) calculates the recall.

$$\text{Accuracy: } \frac{TP + TN}{TP + FP + TN + FN} \tag{3}$$

$$\text{Precision: } \frac{TP}{TP + FP} \tag{4}$$

$$\text{Recall: } \frac{TP}{TP + FN} \tag{5}$$

3 Results and Discussion

In this section, the result from model performance is explained. The result consists of the summary classification model, accuracy, precision, recall, an important feature, and running time. Figure 8 is shown the summary from the decision tree model.

```
=== Summary ===

Correctly Classified Instances      311946               99.9984 %
Incorrectly Classified Instances         5                0.0016 %
Kappa statistic                          0.9997
Mean absolute error                      0
Root mean squared error                  0.0032
Relative absolute error                  0.051  %
Root relative squared error              2.2593 %
Total Number of Instances           311951

=== Confusion Matrix ===

      a      b      c   <-- classified as
 302267      1      1 |      a = BENIGN
      2   5552      0 |      b = FTP-Patator
      0      1   4127 |      c = SSH-Patator
```

Fig. 8. Summary of decision tree model

The first thing, in summary, is the instances that are correctly or incorrectly classified. The decision tree model has classified the instances about 99.9% from all data, and the

incorrect is five instances. Kappa statistic is the calculation of the reliability model. The score of the kappa statistic is 0.99. Mean absolute error is the calculation for measuring the accuracy of a forecasting model because this research not use forecasting. Therefore, the result is zero. Root mean square error is the calculation for evaluating linear model regression. The relative absolute error is the way to measure the performance of a predictive model. The confusion matrix is the table from the prediction result and the true class.

```
 randomForest(formula = Label ~ ., data = trainingData)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 8

        OOB estimate of  error rate: 0%
Confusion matrix:
            BENIGN FTP-Patator SSH-Patator class.error
BENIGN      302269           0           0 0.000000000
FTP-Patator      6        5548           0 0.001080302
SSH-Patator      6           0        4122 0.001453488
```

Fig. 9. Summary of random forest model

Figure 9 illustrates the summary from the random forest model. The random forest model consists of 500 decision trees, and this means there are 500 results to vote in the random forest classification process. Each class's error classification is only 0.001%.

Table 2 represents the accuracy between the decision tree and random forest. Both of them gain 99.9% accuracy. It is a good result for both methods. Decision trees and random forests can classify brute force attacks and normal attacks from network traffic successfully. High accuracy result is the indicator decision tree and random forest successfully recognize brute force attack.

Table 2. Accuracy

Method	Accuracy
Decision tree	99.9%
Random forest	99.9%

Precision gives the information about the positive ratio identifications was correct. Table 3 shows the precision result from all labels in the dataset. The decision tree precision result for normal traffic is 99.9%, which means almost all identification of normal traffic was identified correctly. The decision tree precisions result for brute force attacks are FTP brute force attack 99.9% and SSH brute force attack 99.8%. Almost all results of classification for FTP and SSH brute force attacks were classified correctly. Precision results for the random forest classifier are 100% for normal and FTP brute force attacks, and 99.7% for SSH brute force attacks. Moreover, both decision trees and random forests gain a good precision result.

Table 3. Precision result

Method	Normal	FTP-Patator	SSH-Patator
Decision tree	99.9%	99.9%	99.8%
Random forest	100%	100%	99.7%

Recall function is the indicator of proportion from total actual positives that were identified correctly. Table 4 explained the information about recall results for the decision tree and random forest classifier. Decision tree recall results are 99.9% for normal and FTP brute force attack, 100% for SSH brute force attack. It means almost all normal and brute force attack traffic data was classified correctly. Recall results for random forest are 99.9% for normal traffic and 100% for FTP and SSH brute force attack. Therefore, the random forest can classify almost all proportion normal, FTP, and SSH brute force attacks based on recall results.

Table 4. Recall result

Method	Normal	FTP-Patator	SSH-Patator
Decision tree	99.9%	99.9%	100%
Random forest	99.9%	100%	100%

Table 5 shown the running time decision tree and random forest. Running time is the time it takes to classify network data. On the other hand, the decision trees need 1.61 s to classify network data, while random forests spend 8.68 s.

Table 5. Running time

Method	Running time (s)
Decision tree	1.61
Random forest	8.68

Table 6 explains the important features to define the brute force attack. These important features are important because the network engineer can use these important features to detect brute force attacks in their network environment and disobey other network features. The important features were extracted from the random forest model using the calculation mean decrease in accuracy. Mean decrease in accuracy is the measurement to determine the usefulness of the feature.

Table 6. The important features

Feature	Mean decrease Gini
Destination port	4759.19
Max forward packet length	680.76
Init win bytes backward	664.27
Mean packet length	630.38
Average packet size	611.68

4 Conclusion

The conclusion is that both decision trees and random forest have successfully classified normal and brute force traffic—the result of both classifiers represented by accuracy, precision, and recall results. The accuracy for the decision tree and the random forest is achieved 99.99%. Besides, precision and recall results in normal and brute force attack labels are 99.99%. The classification results of both classifiers have shown similar results. The difference is the calculation of running time. There is a significant difference between decision tree and random forest, and a decision tree only needs 1.61 s to classify network data while random forest spent 8.68 s. The decision tree is faster than the random forest, about 7 s. Therefore, the decision tree is better to use in classifying brute force attacks based on running time. In the future, implement the decision tree in Intrusion Detection System (IDS) and add other brute force attack.

Acknowledgements. The author gives the gratitude to ICEBEHI conference, who gives author the opportunity to publish the manuscript. Thanks to the anonymous reviewers for their precious support to improve the manuscript quality greatly. Thanks to University of New Brunswick who has provided the network attack dataset. They did the great job to makes relevant dataset about network attack. Thanks to Universitas Siber Asia and Institut Teknologi Telkom Purwokerto who has support the author to do this research. Thanks to my colleague who has support the process in writing the paper.

References

1. Imakura A, Sakurai T, Okada Y, Fujii T, Sakamoto T, Abe H (2023) Non-readily identifiable data collaboration analysis for multiple datasets including personal information. Inf Fusion 98:101826. https://doi.org/10.1016/J.INFFUS.2023.101826
2. Belen Saglam R, Nurse JRC, Hodges D (2022) Personal information: perceptions, types and evolution. J Inf Secur Appl 66:103163. https://doi.org/10.1016/J.JISA.2022.103163
3. Joo MH, Kwon HY (2023) Comparison of personal information de-identification policies and laws within the EU, the US, Japan, and South Korea. Gov Inf Q 40:101805. https://doi.org/10.1016/J.GIQ.2023.101805
4. Biju JM, Gopal N, Prakash AJ (2019) Cyber attacks and its different types. Int Res J Eng Technol 4849–4852

5. Sicong Z, Xie X, Xu Y (2020) A brute-force black-box method to attack machine learning-based systems in cybersecurity. IEEE Access 8:128250–128263. https://doi.org/10.1109/ACCESS.2020.3008433
6. Wang D, Ming J, Chen T, Zhang X, Wang C (2018) Cracking IoT device user account via brute-force attack to SMS authentication code. In: RESEC'18, pp 57–60. https://doi.org/10.1145/3203422.3203426
7. Critical M. Brute-force attacks still a major threat. https://www.missioncriticalmagazine.com/articles/93230-brute-force-attacks-still-a-major-threat. Accessed 31 May 2021
8. Verizon (2021) 2021 data breach investigations report. https://doi.org/10.1057/s41280-020-00164-x
9. Muncaster P. Tesco issues 600,000 new clubcards after brute force attack. https://www.infosecurity-magazine.com/news/tesco-issues-clubcards-600k-brute/. Accessed 31 May 2021
10. Georgiev I, Ize T (2018) Arnold: a brute-force production path tracer. ACM Trans Graph 37:1–12
11. Bosnjak L, Sres J, Brumen B (2018) Brute-force and dictionary attack on hashed real-world passwords. In: 2018 41st international convention on information and communication technology, electronics and microelectronics (MIPRO), pp 1161–1166. https://doi.org/10.23919/MIPRO.2018.8400211
12. Otoom AF, Eleisah W, Abdallah EE (2023) Deep learning for accurate detection of brute force attacks on IoT networks. Procedia Comput Sci 220:291–298. https://doi.org/10.1016/J.PROCS.2023.03.038
13. Pérez Gort ML, Olliaro M, Cortesi A (2023) Relational data watermarking resilience to brute force attacks in untrusted environments. Expert Syst Appl 212:118713. https://doi.org/10.1016/J.ESWA.2022.118713
14. Swinhoe D. Brute-force attacks explained, and why they are on the rise. https://www.csoonline.com/article/3563352/brute-force-attacks-explained-and-why-they-are-on-the-rise.html. Accessed 31 May 2021
15. Jo HJ, Yoon JW (2015) A new countermeasure against brute-force attacks that use high performance computers for big data analysis. Int J Distrib Sens Netw 11. https://doi.org/10.1155/2015/406915
16. Amijoyo T, Umar R, Yudhana A (2020) Bruteforce in the hydra and telnet service using the naive Bayes method. J Mantik 4:319–326
17. Mesran, Nasution SD (2020) Peningkatan Keamanan Kriptografi Caesar Cipher dengan Menerapkan Algoritma Kompresi "Stout Codes". J Resti 4:1209–1215
18. Kushwaha V, Maniyath A, Purohit H (2019) Prevention of bruteforce attack using multiple security techniques. Int J Curr Res 11:1910–1912
19. Wanjau SK, Wambugu GM (2021) SSH-brute force attack detection model based on deep learning. Int J Comput Appl Technol Res 10:42–50. https://doi.org/10.7753/IJCATR1001.1008
20. Sandra S, Stiawan D, Heryanto A (2016) Visualisasi Serangan Brute Force Menggunakan Metode K-Means dan Naïve Bayes. Proc Annu Res Sem Proc 2:315–320
21. Nathiya T, Suseendran G (2018) An effective way of cloud intrusion detection system using decision tree, support vector machine and naïve Bayes algorithm. Int J Recent Technol Eng 7:38–43
22. Negandhi P, Trivedi Y, Mangrulkar R (2019) Intrusion detection system using random forest on the NSL-KDD dataset. In: Emerging research in computing, information, communication and applications. Springer Singapore, pp 519–531. https://doi.org/10.1007/978-981-13-6001-5
23. Sharafaldin I, Lashkari AH, Ghorbani AA (2018) Toward generating a new intrusion detection dataset and intrusion traffic characterization. In: ICISSP 2018—proceedings of the 4th international conference on information systems security and privacy, pp 108–116. https://doi.org/10.5220/0006639801080116

24. Zola F, Segurola-Gil L, Bruse JL, Galar M, Orduna-Urrutia R (2022) Network traffic analysis through node behaviour classification: a graph-based approach with temporal dissection and data-level preprocessing. Comput Secur 115:102632. https://doi.org/10.1016/J.COSE.2022.102632
25. Keykhosravi K, Hamednia A, Rastegarfar H, Agrell E (2022) Data preprocessing for machine-learning-based adaptive data center transmission. ICT Express 8:37–43. https://doi.org/10.1016/J.ICTE.2022.02.002
26. Curley C, Krause RM, Feiock R, Hawkins CV (2019) Dealing with missing data: a comparative exploration of approaches using the integrated city sustainability database. Urban Aff Rev 55:591–615. https://doi.org/10.1177/1078087417726394
27. Roshanski I, Kalech M, Rokach L (2023) Automatic feature engineering for learning compact decision trees. Expert Syst Appl 229:120470. https://doi.org/10.1016/J.ESWA.2023.120470
28. Nanfack G, Temple P, Frénay B (2023) Learning customised decision trees for domain-knowledge constraints. Pattern Recogn 142:109610. https://doi.org/10.1016/J.PATCOG.2023.109610
29. Patel HH, Prajapati P (2018) Study and analysis of decision tree based classification algorithms. Int J Comput Sci Eng 6:74–78. https://doi.org/10.26438/ijcse/v6i10.7478
30. Charbuty B, Abdulazeez A (2021) Classification based on decision tree algorithm for machine learning. J Appl Sci Technol Trends 2:20–28. https://doi.org/10.38094/jastt20165
31. Liu R, Li S, Zhang G, Jin W (2021) Depth detection of void defect in sandwich-structured immersed tunnel using elastic wave and decision tree. Constr Build Mater 305:124756. https://doi.org/10.1016/J.CONBUILDMAT.2021.124756
32. Speiser JL, Miller ME, Tooze J, Ip E (2019) A comparison of random forest variable selection methods for classification prediction modeling. Expert Syst Appl 134:93–101. https://doi.org/10.1016/j.eswa.2019.05.028
33. Cheng C, Chen X, De Vos J, Lai X, Witlox F (2019) Applying a random forest method approach to model travel mode choice behavior. Travel Behav Soc. 14:1–10. https://doi.org/10.1016/j.tbs.2018.09.002
34. Couronne R, Probst P, Boulesteix AL (2018) Random forest versus logistic regression: a large-scale benchmark experiment. BMC Bioinform 19. https://doi.org/10.1186/s12859-018-2264-5
35. Peng K, Leung VCM, Zheng L, Wang S, Huang C, Lin T (2018) Intrusion detection system based on decision tree over big data in fog environment. Wirel Commun Mob Comput 2018. https://doi.org/10.1155/2018/4680867

Enhancing COVID-19 Prognostic Accuracy with Machine Learning and Clinical Data Visualization

Phuoc-Hai Huynh[1,2(✉)]

[1] Faculty of Information Technology, An Giang University, An Giang, Vietnam
hphai@agu.edu.vn
[2] Vietnam National University, Ho Chi Minh City, Vietnam

Abstract. Despite the end of the COVID-19 pandemic, its impact on strained hospitals underscores the enduring relevance of timely prognosis in strengthening healthcare systems. This paper proposes to build a predictive framework for COVID-19 patient outcomes through the integration of machine learning algorithms with clinical information. Data were obtained from 2,291 patients at An Giang General Hospital and used to construct and compare six models. Visualization techniques were also applied to identify patterns in clinical features and assess their prognostic relevance. The results indicate that Random Forest and XGBoost consistently achieved superior performance compared to other algorithms. Specifically, Random Forest obtained 95.61% accuracy, 92.24% precision, and an F1 score of 90.04%, while XGBoost reached 95.09% accuracy, 90.51% precision, and an F1 score of 89.90%. The study highlights the advantages of integrating clinical data visualization with machine learning to improve prognostic accuracy, ultimately enhancing healthcare quality and resource management.

Keywords: machine learning · data visualization · prognosis of COVID-19 · clinical data

1 Introduction

In late 2019, COVID-19 was first identified in Wuhan, China. The virus spread quickly to many countries, including Vietnam [1]. It rapidly developed into a global pandemic, leading to millions of cases and deaths. There are more than 82 million people had been infected and at least 1.8 million has died [2]. It created heavy pressure on health systems, especially in countries with poor infrastructure, and led to serious economic and social impacts [3]. The diagnosis of COVID-19 requires the identification of a wide range of symptoms. These manifestations are similar to those of other pulmonary diseases. COVID-19 may worsen into acute pneumonia or respiratory failure, sometimes leading to death. Accurate diagnosis based on clinical presentations is therefore complex and poses significant challenges for patient management and treatment [4].

To address this issue, accurate prediction of a patient's condition plays a crucial role in prognosis and in determining appropriate treatment strategies [5]. Reliable prognostic

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 121–132, 2026.
https://doi.org/10.1007/978-981-95-9233-3_10

assessment enables physicians to develop suitable therapeutic plans. This is particularly essential when the number of hospitalized cases increases. The accurate prognosis supports doctors in identifying high-risk patients. These patients can then be prioritized for intensive care in order to reduce mortality. Moreover, precise prognostic evaluation contributes to minimizing unnecessary costs and preventing the waste of healthcare resources [5].

In this paper, the prognostic model for COVID-19 patients is proposed by utilizing clinical examination data and applying advanced machine learning models for prediction. The experimental results show that machine learning models can accurately predict the treatment outcomes of COVID-19 patients. The dataset is collected from electronic medical records (EMRs) in the Emergency Department of An Giang General Hospital.

To find the best model for accurately assessing patients, our study looked at how well different machine learning models worked. Besides, the study also evaluated the importance of clinical features in influencing prognostic results. These results provide useful information for doctors during treatment.

This research makes several key contributions. First, it enhances existing machine learning models by refining their application in predicting COVID-19 prognosis, thereby increasing their accuracy and reliability. Second, it demonstrates the practical integration of clinical data and machine learning in a real-world healthcare setting, specifically within a Vietnamese hospital context. Finally, it provides a valuable resource for healthcare professionals by identifying the most critical clinical features that influence patient outcomes, aiding in more targeted and efficient treatment strategies.

This work is organized into six parts: introduction, related studies, methodology, model assessment, findings, and conclusion with prospects for further research.

2 Literature Review

Many studies have applied machine learning in medicine. In study [6], the clinical features are proposed for predicting the dangerousness of COVID-19 patients. This study confirmed that machine learning models can effectively and accurately analyze medical data. In addition, Yan et al. used a machine learning model to predict the survival of patients with severe COVID-19 [7]. The findings of this study identified key factors influencing patient treatment outcomes. In another study, natural language processing (NLP) methods were applied to detect important features to predict patients requiring emergency hospitalization [8].

Early prognostic machine learning models have been investigated in studies [9, 10]. These models are capable of providing accurate early predictions to assist physicians in determining appropriate treatment strategies for patients. A study in Middle Eastern hospitals analyzed data from 1,386 patients to predict COVID-19 severity at admission. The results demonstrated high accuracy, ranging from 75.0% to 97% [11]. In study [12], experimental results indicated that age, comorbidities, sex, ventilatory support, and SPO2 were important factors influencing patient mortality.

In summary, although numerous researches have used machine learning models for COVID-19 prognosis. However, the majority have relied on imaging data, PCR testing, or clinical laboratory results [13–16]. There is few studies have utilized initial clinical

examination information, particularly in the context of Vietnam. Therefore, we propose using machine learning algorithms to predict COVID-19 prognosis based on initial clinical examination information at the time of hospital admission. With the current development of hospital information systems in Vietnam, the findings of this research may be integrated to facilitate the early stratification of hospitalized patients, thereby supporting physicians in selecting appropriate treatment strategies.

3 Study Design and Methods

3.1 Data Preparation and Visualization

We implement this study in An Giang Provincial General Hospital. It is one of the An Giang province's leading hospitals and clinical research facilities. The hospital handles an average of 30,000 hospitalizations and 90,000 emergency department visits per year. Over the course of the COVID-19 outbreak, the hospital recorded more than 2,500 patient entries for medical care. This study collected 2,291 patients hospitalized between 2020 and 2022. In the course of data management, we encode patients' sensitive data to safeguard information protection. For each patient, the following variables were examined: sex, age, vital signs, body metrics, and various clinical indicators.

According to the manifestations defined by the WHO, these parameters included 10 components describing the baseline attributes of the patient. They are the causes for admission, the patient's personal medical history, the family medical history, the general physical examination, the respiratory examination, and the circulatory examination. We analyze the dataset using data visualization techniques and the Pearson correlation coefficient.

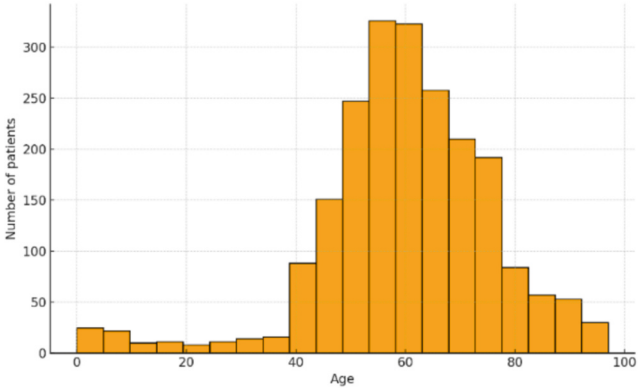


Fig. 1. Patient age distribution

Figure 1 presents the distribution of patient ages, indicating the number of patients in each age group. The age histogram for COVID-19 patients visualizes the age distribution in the dataset. The histogram shows a wide range of ages, from young to old patients. However, the histogram shows that COVID-19 patients in this dataset are primarily

middle-aged, with fewer in the younger and older age groups. Furthermore, the peak around the 60-year mark suggests that this age group is the most affected in this dataset, which could have implications for personalized healthcare.

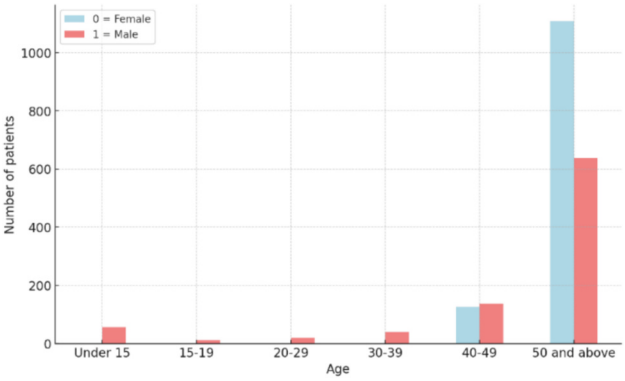


Fig. 2. Sex distribution by age group

Figure 2 illustrates the distribution of patient ages, highlighting the number of patients in each age group. The age histogram for COVID-19 patients visualizes the age distribution in the dataset. The histogram shows a wide range of ages, from young to old patients. However, the histogram shows that COVID-19 patients in this dataset are primarily middle-aged, with fewer in the younger and older age groups. Furthermore, the peak around the 60-year mark suggests that this age group is the most affected in this dataset, which could have implications for personalized healthcare strategies.

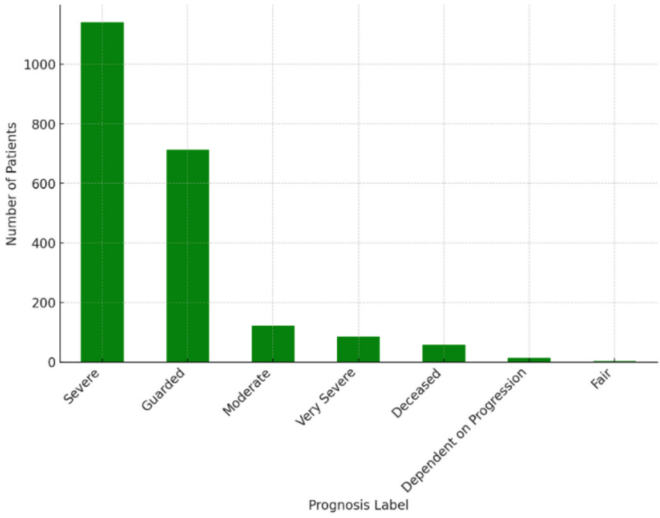


Fig. 3. Distribution of patients across prognosis categories

Figures 3 and 4 show the age distribution by prognosis. The deceased group is predominantly in the "50 and above" age group, with a few in the "40–49" age group. This is dependent on progression and only applies to the "50 and above" age group. In the fair group, there is only the "50 and above" age group. Additionally, most patients in the guarded group are in the "50 and above" age group, with a significant portion in the "40–49" age group and fewer in the younger age groups. The moderate group has only one person aged 40–49. Most are 50 or older. As far as severe status goes, most of the patients are aged 50 and up. There are also a fair number of patients aged 40–49, but not as many younger patients. The majority of people are also 50 or older, while some people in the age range of 40–49 fall into the Very Severe category. The age category "50 and above" has the most patients, despite what the prognosis is. This shows that patients over 65 are more likely to get sick. A lot of individuals in the "40–49" age range are also in the "Guarded," "Severe," and "Very Severe" categories. For younger age groups, all prognosis labels have lower rates. This suggests that younger patients have lower rates of getting worse and dying.

We use the Pearson correlation coefficient Eq. (1) to evaluate the strength of the relationship between age and prognosis. The formula for Pearson's correlation coefficient is as follows:

$$r = \frac{\sum_{i=1}^{n}\left(x_i - \underline{x}\right)\left(y_i - \underline{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \underline{x}\right)^2 \sum_{i=1}^{n}\left(y_i - \underline{y}\right)^2}} \tag{1}$$

where:

x_i *and* y_i represent the individual sample points.

$\underline{x}$ *and* $\underline{y}$ represent the means of the x *and* y samples.

n is the number of sample points.

$\sum$ denotes the summation over all sample points.

The correlation between age and prognosis is about 0.33. The positive value suggests a positive correlation between age and prognosis severity. As people age, the severity of their prognosis tends to increase. The correlation coefficient of 0.33 indicates a moderately positive relationship. It shows that, while age is important in determining prognosis, other factors also play an important role.

Figure 5 shows reveals that males have higher fatal outcomes compared to females. Both genders have similar counts in critical, severe, cautious, fair, moderate, and depending on progress categories. Males have higher counts in more severe prognosis categories (fatal, critical, severe), while females have higher counts in less severe categories (moderate, depending on progress). Males tend to have higher counts in more severe categories, while females have higher counts in less severe categories. In addition, the correlation between sex and prognosis is approximately 0.49. This suggests that the prognosis severity tends to be higher for one sex compared to the other.

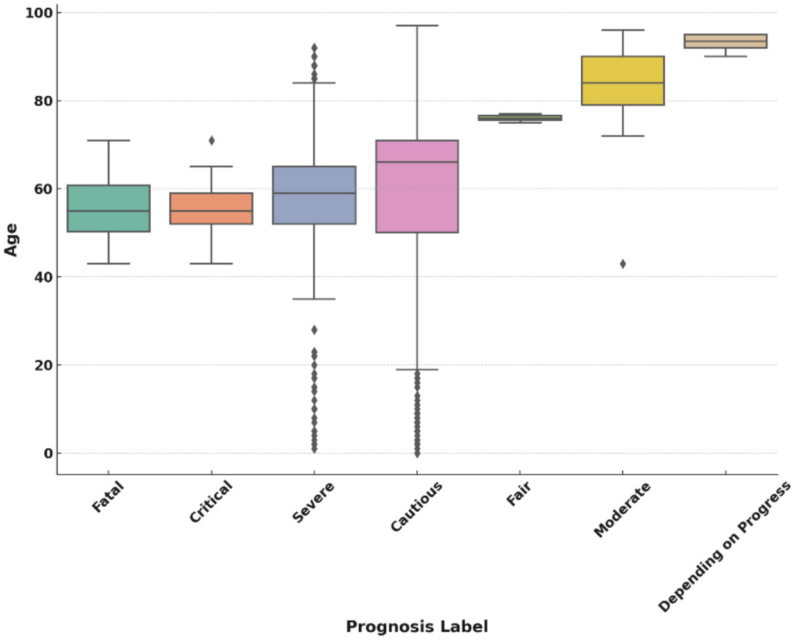


Fig. 4. Age distribution across prognosis categories

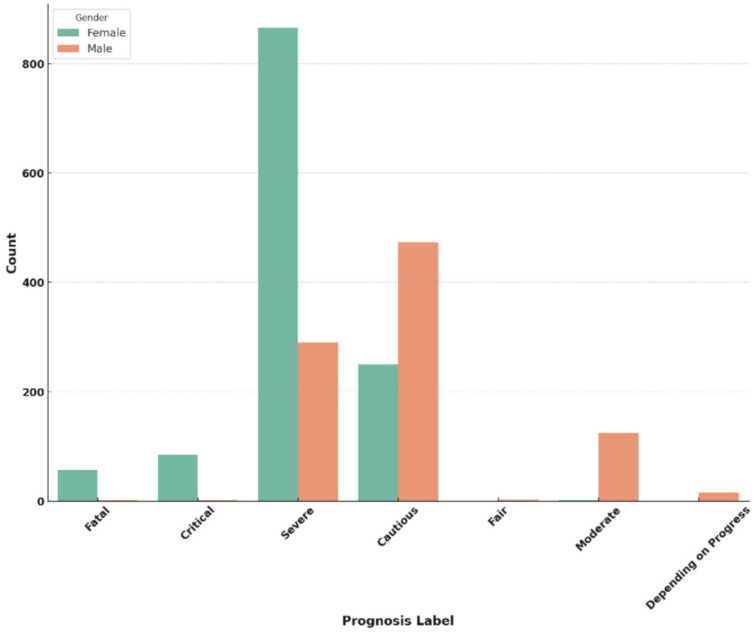


Fig. 5. Prognosis distribution by gender

3.2 Machine Learning-Based Prediction of COVID-19 Patient Prognosis

In this section, we develop prognostic models to identify patients at high risk of adverse outcomes using information available at the time of hospital admission. These models are intended to assist hospitals in stratifying high-risk patient groups for intensive treatment. Classification models are trained on data from 2,291 patients. We use seven models, including *k*NN [17], DT [18], SVM [19], RFs [20], BagDTs [21], and XGBoost [22].

k Nearest Neighbor [17]*:* kNN is a simple yet powerful classification algorithm. It functions by locating the k most similar points in the feature space and generating predictions based on these nearby instances.

Decision Trees: Decision Trees (DT) [18] are constructed using recursive partitioning, where nodes are divided according to a selected splitting rule. In this study, the C4.5 algorithm was applied with the information gain ratio as the splitting criterion to reduce bias.

Support Vector Machines: SVM introduced by Vapnik are based on statistical learning theory and provide a structured approach to classification and regression. SVMs are well known for employing kernel substitution and both SVMs and other kernel-based methods have demonstrated strong practical performance in classification and regression tasks [23, 24].

Random Forest: RF [20] uses the Gini impurity metric to determine feature importance. In this algorithm, each tree in the ensemble is constructed from a sample drawn with a replacement from the training set.

Bagging of Decision Trees: Bagging of Decision Trees (BagDTs) [21] is an ensemble approach that applies bootstrapping to generate multiple subsets from the original dataset. Each subset is created by randomly drawing samples, and individual classifiers are trained in parallel on these subsets. The final prediction is obtained by aggregating the outputs of all classifiers.

Extreme Gradient Boosting (XGBoost): XGBoost [22] is a strong machine learning algorithm that is known for how fast, accurate, and efficient it is. It is a boosting algorithm, a type of ensemble learning that takes the best predictions from several weak learners and puts them all together. This model is versatile in handling a combination of categorical and numeric features. Additionally, XGBoost is particularly effective when only numeric features are present. Its ability to manage numeric data, along with regularization techniques, enhances its overall performance.

3.3 Hyper-Parameter Tuning with Grid-Search-CV

When building machine learning algorithms, selecting suitable hyperparameters is essential as it can greatly influence model performance. In this research, the GridSearchCV procedure was applied to automate the identification of the optimal hyperparameters. This method evaluates every possible parameter configuration, applies cross-validation to assess the model's accuracy under each setting, and then determines the best hyperparameter set. The steps of the GridSearchCV algorithm are outlined as follows:

1. Set a range of values for each parameter.
2. Generate a matrix of all possible parameter value combinations.

Run each combination through an objective function: Each configuration is applied to train and test the model.

3. Select a performance metric: the selected metric may include any measure appropriate for the specific problem, such as accuracy or mean squared error.
4. Evaluate all possible combinations and choose the best one: All parameter sets are compared to determine the best-performing one with the highest performance. The model is then retrained with the optimal hyperparameter setting, and this final model is used for prediction.

We implement the GridSeachCV algorithm in this study to find the optimal hyperparameters of six models. For the SVM parameters, the best gamma is set to "auto" and the best value of C is set to 100. The BagDTs use 100 trees. After tuning the hyperparameters for random forest, we set the maximum depth at 40, the maximum features at "sqrt," the minimum sample split at 2, and the number of estimators at 300. Finally, we use a k-nearest neighbors classifier with a neighborhood of 1 data point. Our k-nearest neighbors model relies on Euclidean distance to identify the nearest neighbors.

4 Evaluation

4.1 Experiments Setup

We implemented six algorithms (kNN, SVM, DT, RF, BagDTs, and XGBoost) for classification tasks. For model training and hyperparameter tuning with GridSearchCV, the Scikit-learn library [25] was utilized. All experiments are implemented on a Linux. The computer consisted of an Intel(R) Core i7-4790 CPU and an 1080Ti GPU.

We use a k-fold cross-validation with $k = 10$, a widely used technique for performance assessment. Classification accuracy serves as the primary metric to quantitatively evaluate model effectiveness, while overall accuracy reflects the efficiency of the classifiers. In addition, the models are assessed using various metrics and run times.

4.2 Classification Results

We are interested in evaluating the prognostic prediction of COVID-19. Table 1 shows a comparison of six models in terms of classification results.

Table 1. Classification results of the models (%)

Models	Accuracy	Precision	Recall	F1 score	Time (s)
k-NN	93.42	86.98	87.44	86.17	0.63
Decision Tree	91.75	83.98	82.98	83.57	0.49
SVM	89.65	81.15	85.88	81.32	2.76

(*continued*)

Table 1. (*continued*)

Models	Accuracy	Precision	Recall	F1 score	Time (s)
RF	**95.61**	**92.24**	89.58	**90.04**	18.76
Bagging	92.87	87.20	86.25	86.09	9.05
XGBoost	*95.09*	*90.51*	**90.38**	*89.90*	12.83

Bold values indicate the best performance for each metric, while italic values denote the second-best results.

Based on the prediction results for COVID-19 patient prognosis, the RF model achieves the highest accuracy (95.61%), followed by XGBoost (95.09%), while the SVM model has the lowest accuracy (89.65%). In terms of precision, the RF model also leads (92.24%), followed by XGBoost (90.51%), with SVM (linear kernel) again at the lowest (81.15%). For recall, XGBoost ranks highest (90.38%), followed by Random Forest (89.58%), and the DT has the lowest recall (82.98%). Regarding the F1 Score, RF tops the list (90.04%), followed by XGBoost (89.90%), with SVM at the bottom (81.32%). In terms of execution time, the DT model is the fastest (0.49 s), followed by kNN (0.63 s), while RF has the longest execution time (18.76 s), followed by XGBoost (12.83 s). In conclusion, for the highest performance without considering execution time, RF is the best choice. For high accuracy with a reasonable execution time, XGBoost is a good option. If execution time is a critical factor, DT and kNN are the best choices with very fast execution times.

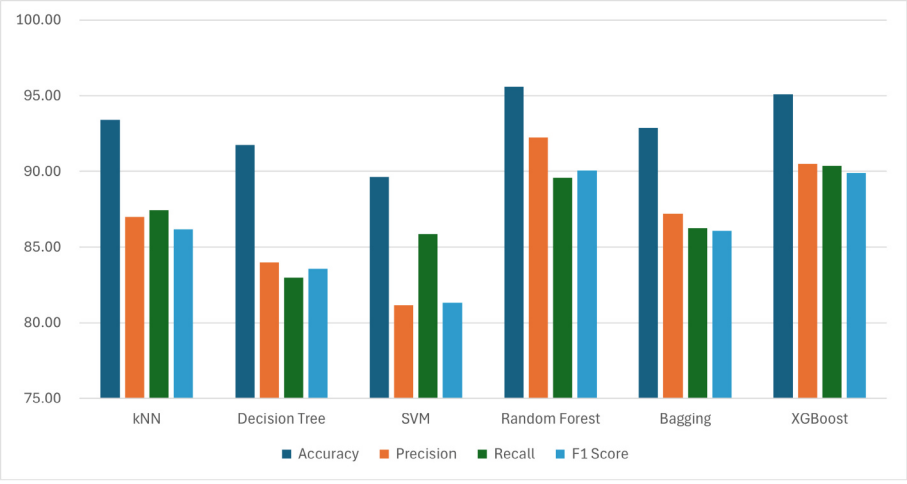


Fig. 6. Model performance metrics comparison

Figure 6 provides a graphical comparison of model performance. The chart presents four evaluation metrics for each model: accuracy, precision, recall, and F1-score. It facilitates a straightforward comparison of performance indicators, highlighting the strengths and limitations of the models.

5 Discussion

We discuss the results in detail, compare them with other studies, highlight the limitations or weaknesses of the study, and present the implications of the study. Based on the tests, the RF and XGBoost are the most accurate at predicting the future, with 95.61% and 95.09% accuracy, respectively. This suggests that machine learning models can be used to accurately predict future outcomes for COVID-19 patients. Using advanced machine learning models, our study revealed similar results to the 2020 Jiang et al. study [6] on COVID-19 severity prediction based on clinical features. While our study derived similar results in terms of prediction accuracy, the authors used EHR data to forecast poor outcomes of patients. Additionally suggesting a prediction models to forecast patient survival with severe COVID-19 was Yan et al. [7]. With several machine learning models, our study revealed either equivalent or better performance than their results.

One problem with the study is that it only looked at how to visualize experimental data and compare and evaluate models. The study primarily focused on comparing and evaluating the performance of different models, without delving deeply into identifying the most critical clinical features that influence prognosis, which could provide more actionable insights for healthcare practitioners. Also, the data came from just one hospital in An Giang, which means that the results might not be applicable to other areas. For this reason, it limits the generalizability of the results to other hospitals or regions with different patient demographics or clinical practices. Additionally, the models were not cross-validated with datasets from other hospitals or regions. Moreover, there may be an imbalance in the distribution of outcomes, such as the number of severe cases versus mild cases, potentially biasing the model's performance toward more common outcomes and reducing accuracy for rarer but clinically significant cases. Lastly, the study did not explore the impact of different feature engineering techniques that could improve model performance or interpretability, leaving room for future research to investigate how various combinations or transformations of features might influence the results.

6 Conclusion and Future Works

In summary, this study demonstrates the effectiveness of machine learning to predict COVID-19 patient outcomes. Using electronic health records, we evaluated several learning algorithms on data from 2,291 patients. The results showed that XGBoost and Random Forest achieved the strongest predictive performance. Data visualization further provided meaningful insights, allowing clear assessment of clinical factors and revealing trends across gender and age features.

In the future, we will aim to improve the explainability of predictive models. Such efforts will advance both the practical application and understanding of these models in healthcare.

Acknowlegement. This research is funded by Vietnam National University Ho Chi Minh City (VNU-HCM) under grant number C2024-16-02.

References

1. Zhu N et al (2020) A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 382. https://doi.org/10.1056/nejmoa2001017
2. Wu YC, Chen CS, Chan YJ (2020) The outbreak of COVID-19: an overview. https://doi.org/10.1097/JCMA.0000000000000270
3. Opinc A, Brzezińska O, Makowska J (2021) The impact of COVID-19 and healthcare system changes on the well-being of rheumatic patients. Reumatologia 59. https://doi.org/10.5114/reum.2021.105424
4. Clinical characteristics of COVID-19 in China. N Engl J Med 382 (2020). https://doi.org/10.1056/nejmc2005203
5. Figliozzi S et al (2020) Predictors of adverse prognosis in COVID-19: a systematic review and meta-analysis. Eur J Clin Invest 50. https://doi.org/10.1111/eci.13362
6. Jiang X et al (2020) Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity. Comput Mater Contin 63:537–551
7. Yan L et al (2020) A machine learning-based model for survival prediction in patients with severe COVID-19 infection
8. Izquierdo et al (2020) Clinical characteristics and prognostic factors for intensive care unit admission of patients with COVID-19: retrospective study using machine learning and natural language processing. J Med Internet Res 22:e21801
9. Booth AL, Abels E, McCaffrey P (2021) Development of a prognostic model for mortality in COVID-19 infection using machine learning. Mod Pathol 34:522–531
10. Chowdhury et al (2021) An early warning tool for predicting mortality risk of COVID-19 patients using machine learning. Cogn Comput 1–16
11. Alrajhi AA et al (2022) Data-driven prediction for COVID-19 severity in hospitalized patients. Int J Environ Res Public Health 19:2958
12. Greco M et al (2022) Outcome prediction during an ICU surge using a purely data-driven approach: a supervised machine learning case-study in critically ill patients from COVID-19 Lombardy outbreak. Int J Med Inform 164:104807
13. Syeda HB et al (2021) Role of machine learning techniques to tackle the COVID-19 crisis: systematic review. JMIR Med Inform 9:e23811
14. Alballa N, Al-Turaiki I (2021) Machine learning approaches in COVID-19 diagnosis, mortality, and severity risk prediction: a review. Inform Med Unlocked 24:100564
15. 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
16. Huynh P-H, Nguyen N-M, Tran T-N, Doan T-N (2023) Improvements in the imbalanced hemogram data classification. In: International conference on electronics, biomedical engineering, and health informatics, pp 347–359
17. Fix E, Hodges JL (1951) Discriminatory analysis. Nonparametric discrimination: small sample performance. Report A 193008
18. Quinlan JR (1986) Induction of decision trees. Mach Learn 1:81–106
19. Vapnik VN, Vapnik V et al (1998) Statistical learning theory
20. Breiman L (2001) Random forests. Mach Learn 45:5–32
21. Breiman L (1996) Bagging predictors. Mach Learn 24:123–140

22. Chen T et al (2015) XGBoost: extreme gradient boosting. R package version 0.4-2, vol 1, pp 1–4
23. Burges CJC (1998) A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov 2:121–167
24. Schölkopf B, Burges CJC, Smola AJ (1999) Advances in kernel methods: support vector learning. MIT Press
25. Pedregosa F et al (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830

Initial Movement of Bionic Foot: Control Mapping Concept

Hartanto Prawibowo[1,2], Rifky Ismail[1,2(✉)], Farika Tono Putri[2,4], Mohammad Tauviqirrahman[1], Robin Novriansyah[2,3], and Sulistia Nur[5]

[1] Mechanical Engineering Department, Universitas Diponegoro, Semarang 50275, Indonesia
rifky_ismail@ft.undip.ac.id
[2] Centre for Bio-Mechanics Bio-Material Bio-Mechatronics and Bio-Signal Processing (CBIOM3S), Semarang, Indonesia
[3] Department of Orthopaedic and Traumatology, Dr. Kariadi Hospital, Semarang, Indonesia
[4] Mechanical Engineering Department, Politeknik Negeri Semarang, Semarang 50275, Indonesia
[5] Nursing Department, Aisyah University, Pringsewu, Lampung, Indonesia

Abstract. Artificial limbs users frequently experience constraints in their range of motion (ROM). This restricted ROM affecting up to 40% of the user's gait cycle. This study presents the development and evaluation of three distinct modes for a bionic leg: Manual mode, automatic mode, and surface Electromyography (sEMG) mode. The methodology in this study involves programming the motor control system, followed by measuring the actual movement of the bionic leg after motor activation. The motor control is divided into three modes: manual control (Manual Mode), surface EMG-based control (EMG Mode), and fully automated control (Automatic Mode). Throughout the experiment, the bionic leg's motion is continuously tracked and recorded using a motion capture camera for detailed analysis. The experimental results reveal that angular settings were varied between 15° and −15°, with dynamic transitions among the modes. Specifically, in manual mode, the angular range averaged between 2° and 18°, with a speed of 50°/s and an acceleration of 200°/s^2. In sEMG mode, the angular range was between 5° and −13°, with a speed of 25°/s and an acceleration of 100°/s^2. Automatic mode achieved an angular range of 17° to −20°, with a speed of 10°/s and an acceleration of 80°/s^2. Future advancements will focus on refining these modes to more accurately replicate human gait, including the integration of dorsiflexion and plantarflexion angles to align with natural walking patterns.

Keywords: Bionic · Legs · Foot · sEMG · Angle Sensor · BLDC

1 Introduction

The problem of disability has emerged as a crucial issue in Indonesia in recent years. According to data from the Coordinating Ministry for Human Development and Culture of the Republic of Indonesia, in 2021, there were 22.97 million people with disabilities, which included 8.5% of the total population of Indonesia. Besides being caused by

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 133–145, 2026.
https://doi.org/10.1007/978-981-95-9233-3_11

morbidity in certain diseases, disabilities are also caused by many driving accidents, whose trends increase every year.

This condition shows the need for serious attention and actions from various parties to reduce the number of traffic accidents and provide better support for people with disabilities.

The number of artificial foot users in Indonesia has increased every year. Data collected from the Central Statistics Agency shows that as many as 568,020 people with disabilities have difficulty walking or climbing stairs. This increase is also in line with the increase in the number of imports of medical devices every year.

Persons with disabilities in Indonesia still use mechanical artificial feet, known as Solid Ankle Cushioned Heels (SACH), and Energy Storage and Return (ESAR). However, the two products available on the market cannot move rotating on the ankle joint. As a result, there is a tendency for voltage concentration at a pounding point in the ankle joint, which can cause discomfort and other biomechanical problems [21, 22].

The bionic foot is a substitute for the lower limbs by integrating the dorsiflex and plantarflexion movements in the ankle joint, which aligns with the current cycle. Some previous studies show that amputation patients often experience limitations in their walking patterns. However, with the use of bionic feet, motion patterns can be improved because of the ability of ankle joints to move again after the installation of bionic feet [1, 2].

SACH foot users are limited to motion limitations of up to 40% of the current cycle, especially in the dorsiflexion and plantarflexion movements. This limitation makes conventional artificial legs unable to mimic the kinematics and kinetics of the gait cycle (Fig. 1) [3].

Previous research has investigated bionic feet' design and manufacturing process and carried out a series of motion tests. The results allow bionic feet to be tested in humans in the next stage of research [4, 5].

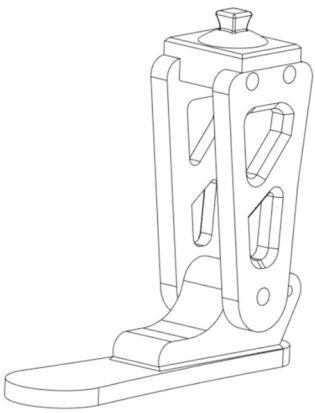


Fig. 1. Bionic foot from previous model [4]

The use of bionic feet with EMG sensors and angular sensors becomes the best alternative to prevent motor excess motion. It can avoid mechanical damage that is

caused by errors in the process of dorsiflexion and plantarflexion. The sEMG method, by attaching muscle sensors to the patient's muscles, can move the motor continuously. And there is no mistake in the process of running. The angle sensor used can limit and adjust the mechanical angle of the frame if the desired angle has not been rejected. Angle angles in the right frame can also reduce excessive design, such as stopper design or other stopper designs [23–25].

Dual sEMG Mode and continuous angle sensors can affect the results of the reading that enters the system. The mechanical system that is carried out on the angle sensor and the muscle signal system released by the seemed can make it easier for patients to do dorsiflexion and plantarflexion movements. So that it can replace the style of walking in a normal person [6, 7].

The challenges in developing a bionic leg primarily arise from the complexity of translating intended movements into the bionic leg's functionality. By refining the mapping process, we aim to streamline previous research efforts. This mapping will specifically focus on essential movement patterns necessary for the running cycle. Moreover, it promises to simplify the design, manufacturing, and mechanical or clinical testing phases [3, 8–11].

2 Materials

2.1 EMG Sensor

Surface electromyography (SEMG) as shown in Fig. 2 is a non-invasive method for evaluating muscle function by recording muscle activity from the skin's surface. This technique has various applications in various contexts, including clinical diagnosis, neurological/neuromuscular problems, laboratory research on the current cycle, and research experiments on human muscles [12–15].

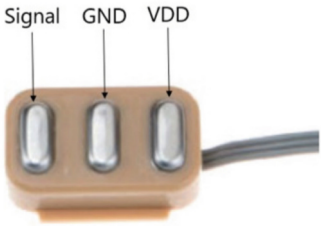


Fig. 2. Surface EMG

The electric potential generated during muscle contraction is captured and analyzed through a computing system, frequently utilizing artificial intelligence (AI). Key components of this system include signal detection, where surface electromyography (sEMG) signals are acquired via non-invasive electrodes. The recorded signals yield valuable data on the timing and intensity of superficial muscle activity.

2.2 Angular Sensor

Angle measurement utilizing magnetic encoder AS5600 (Fig. 3), which is used to determine how much of the angle at the frame is driven by the motor. The magnetic encoder works on the frame system with a magnet in the shaft to read the magnitude. This sensor will send an angular quantity to the system in units of angle degrees. To prevent unreadable results from other sensors, a magnetic encoder is used to avoid errors in reading angles that can be fatal from the motor that will not stop at the right time. A small and concise shape can help in the design process for magnetic encoders [16].

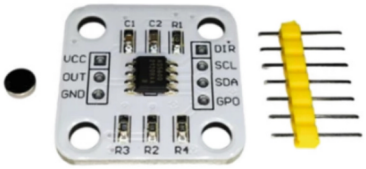


Fig. 3. Magnetic encoder AS5600 as angle measurement

2.3 Linkage

Linkage serves to connect between ankle joint, bracket, and linear screw. Linkage design in this current study is presented in Fig. 4. Where the movement of the Liner Screw will affect the length of the bracket and the angle needed to carry out the process on the person running normally, optimizing the linkage itself to find out with the existing motor can reach the speed of motion at the angle that is appropriate to the person running normally (the gait cycle).

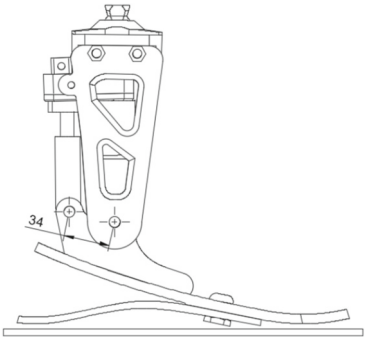


Fig. 4. Present model

A previous design study shows a linkage length of 34 mm from the center of the linkage to the center of the ankle. The angle requirement needed to move the bionic V1 foot is 18.1°. The number 18.1° is obtained from the initial study of the bionic foot of V1.

2.4 Microcontroller

EMG Surface Active Mode in the remote interface is transferred and diverted into EMG mode, then contracted on the patient's leg muscles. Amputation: After systems receive muscle contraction stimulation, the signal will be an amplifier that produces a signal that tends to be clean and stable from EMG. The signal is then inputted with Magnetic Rotary Encoder AS5600 and then sent together with *Arduino Promini* to send a signal to the ESC. ESC gets a signal to set the amount of power incoming, and after moving the BLDC motor, then the Motor BLDC will move the gearbox to be transmitted to the transmission circle, and then the environment will raise the frame mechanism in the bionic feet. Mechanically, the frame system will move the magnetic encoder, which is installed with the frame on the frame on the joint [17].

2.5 Video Capture Camera

Video recording camera processing is used to capture angular data, facilitating the measurement process. Video processing results can be seen in Fig. 5. This data is employed to assess the angular magnitude, angular velocity, and angular acceleration. The system utilizes a motion camera, commonly used for recording walking movements. The angles captured will be automatically detected by the software and processed through an image generator for detailed analysis. By collecting data from individuals walking for a continuous duration of three minutes, we aim to capture a representative sample of normal gait patterns. This data is then processed and analyzed using a motion capture system, which provides detailed insights into the biomechanics of walking.

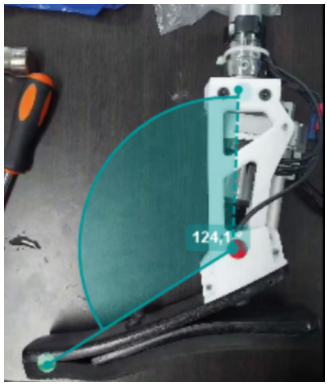


Fig. 5. Video capture camera

3 Method

3.1 Manual Mode

After the component was made and component assembly was carried out, trials were carried out regarding the system that had been built, starting from manual mode, namely by moving the dorsiflexion and plantarflexion angles by pressing the button continuously.

The manual model is carried out to determine whether the system can work properly in the mechanical system. If a mismatch is found in the frame that collides or the anchor angle is not achieved, then the mechanical system can be replaced. Starting from the manual setting, the aim is to set the range of motion (ROM) on the dorsiflexion and plantarflexion angles. Next, the angular swing speed is set to produce the desired speed and acceleration from the bionic leg.

3.2 Automatic Dorsiflexion and Plantarflexion Mode

In automatic movement mode, the dorsiflexion and plantarflexion angles will carry out the gait cycle phase automatically without any external intervention by setting the position and speed first to limit dorsiflexion and plantarflexion movements. The limitation aims to prevent damage to the mechanical system caused by the high angle worked by the motor. The motor is limited by the angle read from the magnetic encoder in the system. Then, the sensor will send the amount of angle that has been achieved by the mechanical system into the control.

3.3 Surface Electromyography (EMG) Mode

sEMG Mode is a mode in which muscle contraction data is collected from the sEMG sensor. Make settings in the "setting" menu by changing the RoM first, then setting the dorsiflexion and plantarflexion angle speed in the gait setting menu and acceleration in the same menu need to set the PID in the "PID Setting" menu to adjust KP, KI, KD, namely the motor power. Next, the muscle will be contracted to activate the sEMG sensor. Then, the sEMG sensor and angle sensor will be adjusted to the control to move the motor. With power from the ESC, the BLDC motor will rotate and stop according to commands from the control. Next, the frame will move in a swing and be limited to the previously set limits.

3.4 Interface

The interface (Fig. 6) was created to display simple text to test the bionic foot system that has been created. In the process, it is classified into six settings and movement modes, "Main" being the first mode, namely the dorsiflexion and plantarflexion movement modes. Whose function is to test the speed of the motion cycle continuously. "Manual" mode is a mode where dorsiflexion and plantarflexion movements are done manually using buttons. "Settings", "Gait Settings", and "PID Settings" are tool settings for the system. "EMG Mode" is a dorsiflexion and plantarflexion movement mode that automatically uses muscle contractions using surface electromyography [18–20].

Fig. 6. Bionic foot interface

3.5 Measurement

Extensive measurements across bionic foot operational modes can be crucial to be conducted in order to thoroughly understand and enhance the performance of the bionic foot. This research is focused on evaluating the bionic leg's performance across three distinct modes: Manual, Automatic, and sEMG. Each mode presents unique characteristics and control mechanisms, necessitating a comprehensive assessment to determine the operational benefits and limitations of the system.

The evaluation involved sequential testing of the bionic leg in each mode. The objective is to measure how effectively and efficiently each mode replicates human gait patterns. Specifically, the analysis targets dorsiflexion and plantarflexion angles, which are critical for mimicking natural walking movements. The angles examined are 15° for dorsiflexion and −15° for plantarflexion, with both movements integral to simulating a walking motion.

This approach will yield valuable data regarding the performance of the bionic leg in each mode and its contribution to the overall system functionality. A detailed analysis of these measurements is anticipated to offer deeper insights into how each mode influences user dynamics and comfort, and will serve as a foundation for further advancements and refinements in bionic leg design.

4 Result

4.1 Angles

The results from the camera capture indicate that movements recorded from various angles and under different control settings yield different results. These variations in movement could be attributed to previously configured parameters, but due to the inherent limitations of the tool, the motor may produce varying outcomes.

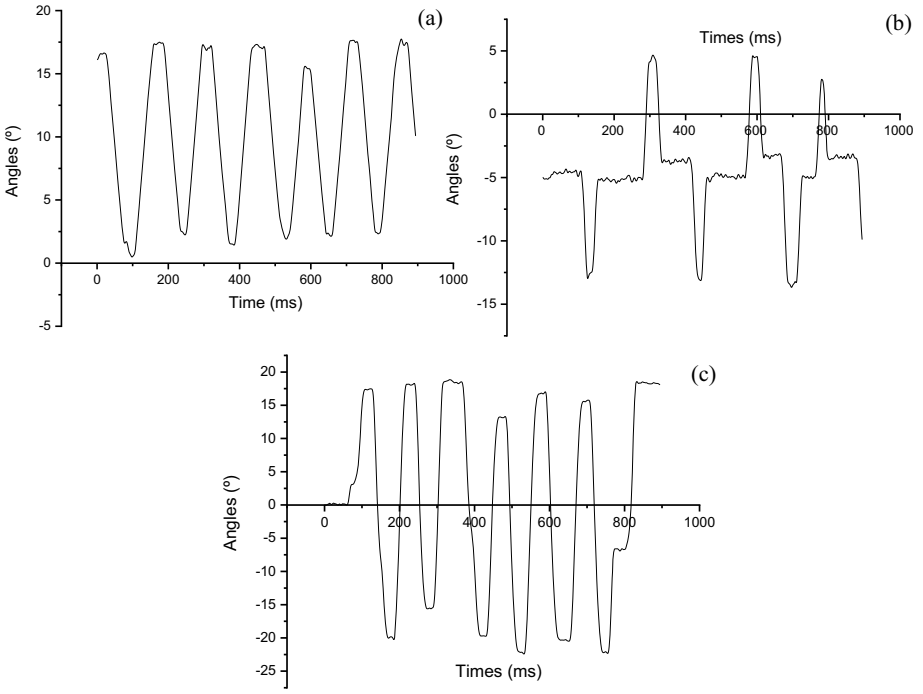


Fig. 7. The angle determined in manual mode (a), the angle recorded during EMG mode (b), and the angle quantified in automatic mode (c)

The angles shown in Fig. 7.A show that in manual mode, the angle can only be formed in the dorsiflexion position. Conversely, in the plantar flexion position, the manual mode cannot accommodate the movement. This is because the manual mode operates at high speed and only produces angles in the range of 2°–18°, all of which are included in the dorsiflexion category.

In Fig. 7.B, the observed angles indicate the EMG mode, where there is noise interference in the angular movement of the bionic leg. This interference is caused by the inaccuracy of the EMG signal reading on the bionic leg due to muscle contraction. However, the EMG mode is able to accommodate dorsiflexion and plantarflexion angles, with an angle range between 5° and −13°. The EMG mode has been effective in interpreting the movements required for walking activities.

In Fig. 7.C, the angles shown indicate manual mode, also known as continuous motion mode. In this mode, the motion is performed continuously with the angles being adjusted continuously. The resulting angles, ranging from 17° to −20°, are able to accommodate both dorsiflexion and plantarflexion in walking.

4.2 Angular Velocity

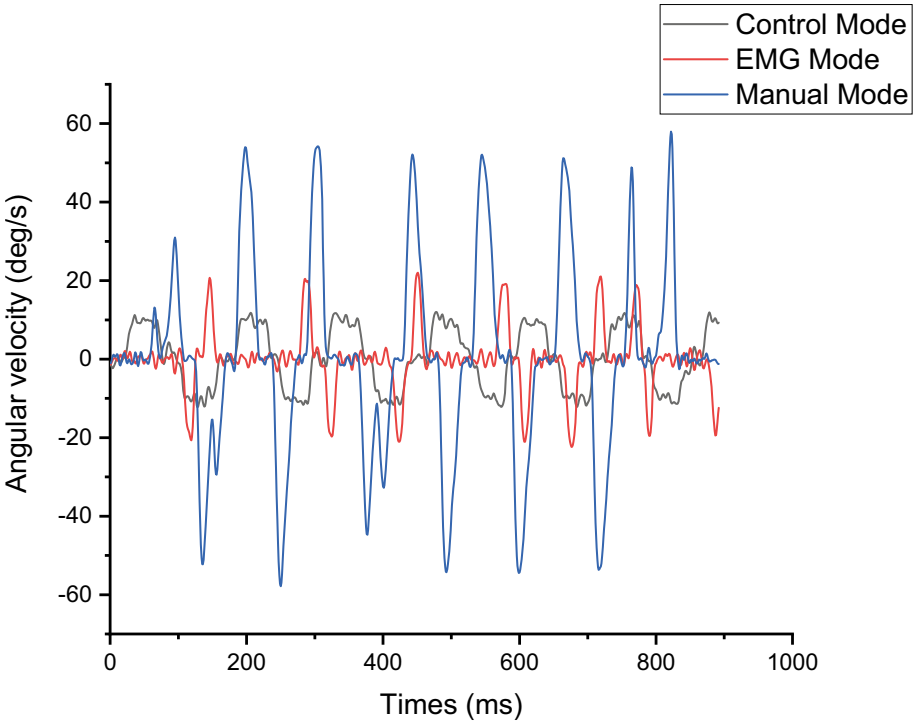


Fig. 8. Angular velocity graph

The angular velocity in both modes shows significant differences. Illustrates that the velocity in these modes varies greatly (Fig. 8). In manual mode, the highest angular velocity reaches a maximum of 50°/s, due to the absence of restrictions on the initial and final velocities. In contrast, in automatic mode, the angular velocity is relatively lower, around 10°/s, because the angular magnitude and angular velocity have been previously set. While in EMG mode, the average velocity is 25°/s, which is due to the limitation of muscle contraction regulated by the EMG signal.

4.3 Angular Acceleration

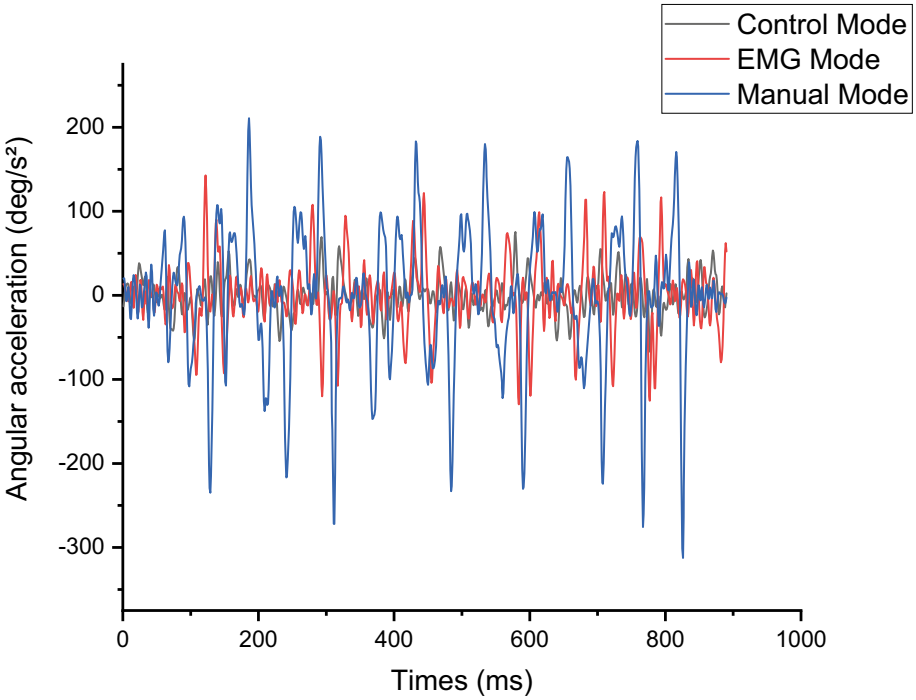


Fig. 9. Angular acceleration graph

Angular acceleration across all modes results from high speeds during significant angular movements. In manual mode, the acceleration peaks with an average of up to 200°/s^2. Conversely, in EMG mode, the average acceleration is up to 100°/s^2. In control mode, however, the average acceleration is the lowest, at 80°/s^2. This acceleration evolves in conjunction with angular velocity over time (Fig. 9). High levels of acceleration are crucial for assessing the impact of dorsiflexion and plantarflexion on human gait.

5 Discussion

The study's findings detail the angular measurements for the three modes under investigation. Prior to testing all modes, angular settings were configured from 15° to −15°, and the modes transitioned actively throughout the process. In manual mode, the average angular range was recorded between 2° and 18°, with a speed of 50°/s and an acceleration of 200°/s^2. For the EMG mode, the average angular range was between 5° and −13°, with a speed of 25°/s and an acceleration of 100°/s^2. In automatic mode, the angular range achieved extended from 17° to −20°, with a speed of 10°/s and an acceleration of 80°/s^2.

These results reveal that each mode has distinct limitations. The manual mode, despite demonstrating high speed and acceleration, fails to achieve substantial overall speed. Conversely, the automatic mode meets the angular range necessary for simulating walking motion, but does not match the speed and acceleration of human walking. The EMG mode, currently under development, necessitates additional research into muscular dynamics to enhance its performance. Further research strongly advocates for the incorporation of an Inertial Measurement Unit (IMU). According to Jaehuu's study, there is an intention to integrate an IMU system into the bionic leg, enhancing its functionality and performance [14].

6 Conclusion

In this study, the bionic leg is anticipated to achieve gait cycles with high efficiency and precision. Before testing, the angular settings were adjusted from 15° to −15°, with modes transitioning dynamically throughout the process. In manual mode, the angular range averaged between 2° and 18°, with a speed of 50°/s and an acceleration of 200°/s^2. In EMG mode, the average angular range was between 5° and −13°, with a speed of 25°/s and an acceleration of 100°/s^2. In automatic mode, the angular range achieved was from 17° to −20°, with a speed of 10°/s and an acceleration of 80°/s^2.

Further research is essential for the next phases of development. The IMU (Inertial Measurement Unit) proves to be an ideal measurement tool for simplifying the process of measuring and providing feedback on the motor. Conversely, system upgrades and improved coding can facilitate the feedback process by using magnetic encoders as sensors, with EMG serving as the primary actuator trigger.

References

1. Lura DJ, Wernke MW, Carey SL, Kahle JT, Miro RM, Highsmith MJ (2017) Crossover study of amputee stair ascent and descent biomechanics using Genium and C-Leg prostheses with comparison to non-amputee control. Gait Post 58:103–107. https://doi.org/10.1016/j.gaitpost.2017.07.114
2. Masum H, Bhaumik S, Ray R (2014) Conceptual design of a powered ankle-foot prosthesis for walking with inversion and eversion. Procedia Technol 14:228–235. https://doi.org/10.1016/j.protcy.2014.08.030
3. Cheng S, Bolivar-Nieto E, Gregg RD (2021) Real-time activity recognition with instantaneous characteristic features of thigh kinematics. IEEE Trans Neural Syst Rehabil Eng Publ IEEE Eng Med Biol Soc 29:1827–1837. https://doi.org/10.1109/TNSRE.2021.3107780
4. Ismawan AR, Ismail R, Novriansyah R, Setiyana B, Ariyanto M, Prahasto T (2021) Conceptual design of bionic foot for transtibial prosthesis. In: 2021 IEEE international biomedical instrumentation and technology conference (IBITeC), pp 124–129. https://doi.org/10.1109/IBITeC53045.2021.9649405
5. Prawibowo H, Putri FT, Ismail R, Tauviqirrahman M, Novriansyah R, Setyoadi Y (2023) Finite element analysis on a bionic foot prosthesis model during walking gait phases. In: 2023 IEEE international biomedical instrumentation and technology conference (IBITeC), pp 135–140. https://doi.org/10.1109/IBITeC59006.2023.10390945
6. Putri FT, Caesarendra W, Kr G, Glowacz A (2023) Human walking gait classification utilizing an artificial neural network for the ergonomics study of lower limb prosthetics, pp 647–665

7. Hussain T, Iqbal N, Maqbool HF, Khan M, Awad MI, Dehghani-Sanij AA (2020) Intent based recognition of walking and ramp activities for amputee using sEMG based lower limb prostheses. Biocybern Biomed Eng 40:1110–1123. https://doi.org/10.1016/j.bbe.2020.05.010
8. Au SK, Weber J, Herr H (2009) Powered ankle-foot prosthesis improves walking metabolic economy. IEEE Trans Robot 25:51–66. https://doi.org/10.1109/TRO.2008.2008747
9. Herr HM, Grabowski AM (2012) Bionic ankle-foot prosthesis normalizes walking gait for persons with leg amputation. Proc R Soc B Biol Sci 279:457–464. https://doi.org/10.1098/rspb.2011.1194
10. Song H et al (2024) Continuous neural control of a bionic limb restores biomimetic gait after amputation. Nat Med. https://doi.org/10.1038/s41591-024-02994-9
11. Hill D, Herr H (2013) Effects of a powered ankle-foot prosthesis on kinetic loading of the contralateral limb: a case series. In: IEEE international conference on rehabilitation robotics. https://doi.org/10.1109/ICORR.2013.6650375
12. Lalumiere M, Villeneuve C, Bellavance C, Goyette M, Bourbonnais D (2020) Patterns of lower limb muscular activity and joint moments during directional efforts using a static dynamometer. BMC Biomed Eng 2:1–12. https://doi.org/10.1186/s42490-019-0035-7
13. Mohiuddin et al (2021) Mater Des 11:5035–5040
14. Ryu J, Lee BH, Maeng J, Kim DH (2019) SEMG-signal and IMU sensor-based gait sub-phase detection and prediction using a user-adaptive classifier. Med Eng Phys 69:50–57. https://doi.org/10.1016/j.medengphy.2019.05.006
15. Park J et al (2023) Imperceptive and reusable dermal surface EMG for lower extremity neuroprosthetic control and clinical assessment. npj Flex Electron 7:1–11. https://doi.org/10.1038/s41528-023-00282-z
16. Chinimilli PT, Redkar S, Sugar T (2019) A two-dimensional feature space-based approach for human locomotion recognition. IEEE Sens J 19:4271–4282. https://doi.org/10.1109/JSEN.2019.2895289
17. Latif T, Ellahi CM, Choudhury TA, Rabbani KS (2008) Design of a cost-effective EMG driven bionic leg. In: Proceedings of ICECE 2008—5th international conference on electrical and computer engineering, pp 80–85. https://doi.org/10.1109/ICECE.2008.4769177
18. Viteckova S, Kutilek P, Jirina M (2013) Wearable lower limb robotics: a review. Biocybern Biomed Eng 33:96–105. https://doi.org/10.1016/j.bbe.2013.03.005
19. Kang M, Shin H, Cho Y, Park J, Nagwade P, Lee S (2022) Triboelectric neurostimulator for physiological modulation of leg muscle. Nano Energy 103:107861. https://doi.org/10.1016/j.nanoen.2022.107861
20. Wang K, Tobajas PT, Liu J, Geng T, Qian Z, Ren L (2019) Towards a 3D passive dynamic walker to study ankle and toe functions during walking motion. Robot Auton Syst 115:49–60. https://doi.org/10.1016/j.robot.2019.02.010
21. Van TLE (2017) Evaluation of trans-femoral prosthesis function using finite element analysis, Sept 2017. [Online]. Available: https://shibaura.repo.nii.ac.jp/?action=repository_action_common_download%5C&item_id=97%5C&item_no=1%5C&attribute_id=20%5C&file_no=1
22. Delussu AS, Paradisi F, Brunelli S, Pellegrini R, Zenardi D, Traballesi M (2016) Comparison between SACH foot and a new multiaxial prosthetic foot during walking in hypomobile transtibial amputees: physiological responses and functional assessment. Eur J Phys Rehabil Med 52(3):304–309
23. Hu B, Rouse E, Hargrove L (2018) Fusion of bilateral lower-limb neuromechanical signals improves prediction of locomotor activities. Front Robot AI 5:78. https://doi.org/10.3389/frobt.2018.00078
24. Spanias JA, Perreault EJ, Hargrove LJ (2016) Detection of and compensation for EMG disturbances for powered lower limb prosthesis control. IEEE Trans Neural Syst Rehabil Eng 24(2):226–234. https://doi.org/10.1109/TNSRE.2015.2413393

25. Spanias JA, Simon AM, Ingraham KA, Hargrove LJ (2015) Effect of additional mechanical sensor data on an EMG-based pattern recognition system for a powered leg prosthesis. In: International IEEE/EMBS conference on neural engineering NER, vol 2015, July 2015, pp 639–642. https://doi.org/10.1109/NER.2015.7146704

Classification of Stress Types Based on Electrodermal Activity Signal Using Multi-layer Perceptron Method

Clara Tyas Ratri[1], Osmalina Nur Rahma[1,3(✉)], Sayyidul Istighfar Ittaqillah[1,2], Rima Firda Alfisyah[1], Atin Asna Octavia[1,2], Fadli Ama[1,3], Khusnul Ain[1,3], and Rifai Chai[1,4]

[1] Biomedical Engineering Study Program, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia
osmalina.n.rahma@fst.unair.ac.id

[2] Biomedical Engineering Master Study Program, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia

[3] Biomedical Engineering Innovation Study Group, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia

[4] Department of Telecommunication, Electrical, Robotics and Biomedical Engineering, Swinburne University of Technology, Victoria, Australia

Abstract. In recent decades, mental health has become a concern due to the prevalence of chronic stress, which can lead to various health problems. One promising non-invasive method that has been widely used to detect stress is electro-dermal activity (EDA). EDA is considered an accurate and sensitive stress measurement method that analyzes skin conductance (SC) due to changes in sympathetic nerve activity. The SC signal consists of phasic and tonic com-ponents, so it is necessary to deconvolve it when analyzing stress conditions. In this study, a convex optimization method (cvxEDA) was proposed to de-compose and analyze EDA signals to assess the emotional state of chronic stress sufferers. This method can help identify different types of stress, such as physical, cognitive, or emotional stress, and aid in handling them according to psychological conditions. Stress-type classification will be performed using a Multi-Layer Perceptron (MLP) algorithm. The study involved 20 subjects who underwent 3 measurement sessions with different stress stimuli to increase their stress levels. The study results demonstrated that cvxEDA method successfully identified and separated the SC components. The phasic component (SCR) was found to describe the subjects' physiological responses to the stressors. The MLP model was used to classify 3 different stress types and achieved better accuracy value in the classification of cognitive stress with an accuracy of 82% and an F1-score of 80.13% for cognitive stress. This research provides implications in the form of the potential for early detection of stressful situations so that it can provide treatment according to the stress-type.

Keywords: stress types · electrodermal activity · multi-layer perceptron · cvxEDA

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 146–159, 2026.
https://doi.org/10.1007/978-981-95-9233-3_12

1 Introduction

Mental health has grown in importance and is now a common conversation subject in recent decades. A person's emotional, psychological, and social well-being in managing stress and adjusting to life's challenges is known as their mental health [1, 2]. Problems in mental health can lead to mental disorders. According to the WHO, 1 in 8 people worldwide suffer from mental illnesses in 2019. Stress is one of the many diverse forms of mental diseases that are characterized by clinically significant changes in cognition, emotion control, or behaviour [3]. The World Health Organization (WHO) has classified stress as a global pandemic because of the severe consequences it causes, which include higher rates of mental illness and suicide as well as increased morbidity and mortality [4]. Based on this description, stress is one of the factors that cause mental illness or psychiatric disorders.

Stressful conditions can be characterized by high arousal values which is a symptom due to an increase in a person's physiological and psychological activity when the adrenal glands produce adrenaline due to increased sympathetic nervous system activity [5]. Increased arousal is indicated by increased sympathetic nervous system activity and HPA activation. The HPA axis is the body's neuroendocrine system that regulates reactions to stress and maintains the body's physiological homeostasis [6]. The definition of stress based on WHO itself is the body's reaction or response to a pressure or psychosocial stress [1]. In general, stress based on the stressor can be categorized into physical stress and psychological stress [7]. Physical stress can occur due to a stimulus that burdens the physical body, causing tension such as physical trauma and has a relationship with physiological factors such as genetic factors [8, 9]. Meanwhile, psychological stress is stress due to events that affect psychological factors and can be classified into 2 types, namely emotional and cognitive stress [10].

Prolonged physical and psychological stress can affect all body systems including the musculoskeletal, respiratory, cardiovascular, endocrine, digestive, nervous, and reproductive systems [11]. The effects of stress on systems in the body can potentially lead to the emergence of other diseases. Therefore, it is necessary to monitor a person to determine the type of stress experienced as an effort to prevent the emergence of diseases due to prolonged stress. One promising non-invasive method that has been widely used to detect stress is electrodermal activity (EDA). The use of EDA is often used and applied in the field of psychology, one of which is to determine the effects of stress on body functions related to changes in sweat gland activity that reflect the intensity of the body's arousal value [12, 13]. The EDA signal results consist of Skin Conductance Level (SCL) which represents the tonic component and Skin Conductance Response (SCR) which represents the phasic component. The overlapping phenomenon between the two signals needs to be overcome by separating the tonic and phasic components using deconvolution [14, 15].

In [16], a performance comparison of several deconvolution methods was carried out including the Convex Optimization (cvxEDA), Sparse Deconvolution (SparsEDA), Continuous Decomposition Analysis (CDA), Discrete Decomposition Analysis (DDA) and Dynamic Causal Modeling (DCM) methods. Based on this research, cvxEDA showed good results in decomposing EDA signals into tonic (SCL) and phasic (SCR) components. Based on research [17] the cvxEDA method has been applied as a pre-processing

step to remove artifacts from each EDA signal and at the same time extract its main components: tonic and phasic components. In these studies, only signal processing was carried out to get the best EDA signal quality so that it was not yet at the EDA signal classification stage. Research related to stress detection is carried out in research [18] where this study uses Super Vector Machine (SVM) and Artificial Neural Network (ANN) as EDA signal classifiers combined with Electrocardiograph (ECG) in car drivers. The results of this study show that accuracy with ANN is better than SVM. The research [19] also utilizes the Galvanic Skin Response (GSR) value which is one type of EDA where this signal will be combined with Photoplethysmograph (PPG) as input for classification using Linear Discriminant Analysis (LDA). In this study, a good value was obtained which means that the conductivity signal on the skin can provide good information on the state of stress. However, the research only focuses on stress classification in general.

Classification of EDA signals has been carried out in research [20] with various methods to classify stress with different stressors. The variations of methods used are Support Vector Machines (SVM), K-Nearest Neighbors (KNN) and Multi-Layer Perceptron (MLP). Based on this research, it is known that the best classification performance is obtained using the Multi-Layer Perceptron (MLP) method which has higher accuracy than the other two methods. The use of MLP also provides good value in research [21] by using several sensors, one of which is EDA but in this study, it does not go through the deconvolution stage so that the EDA signal used contains phasic and tonic components respectively. The combination of deconvolution with classification has been done in previous studies. In research [22] has done a combination of cvxEDA deconvolution with SVM classification. This study provides good value on classification between stress and non-stress but not good enough on multiclass classification, namely 4 types of stress based on stressors. In research [23], a good combination of cvxEDA method and Extreme Learning Machine (ELM) classification has also been done but the classification is only limited to 2 conditions: cognitive and emotional stress.

Based on the previous research, there is a research gap in the absence of a combination of the EDA signal deconvolution method with the classification method to classify stress based on its stressors into physical, cognitive, and emotional stress. The gap in the use of the best deconvolution and classification methods in previous studies is the potential proposed in this study. In this study, a combination of the two methods is carried out, namely the deconvolution method with cvxEDA and the classification method with MLP. This research focuses on using a combination of the cvxEDA method with MLP so that it is expected to provide better deconvolution results in order to get more optimal classification results. This research was conducted to determine the type and situation of stress as an effort to reduce the potential for disease due to prolonged stress. The classification of stress types is also carried out to find out what stressful situations are being experienced by a person, so that they can be conditioned with treatments according to the type of stress.

2 Material and Methods

Figure 1 shows the research procedures in this research that conducted in 5 steps. The data is collected from the public dataset and then the data will be preprocessed with deconvolution using cvxEDA. After preprocessing data, the data will be extracted features to get 7 statistical features. After all the features are obtained, the classification will be carried out using MLP classifier and will be evaluated with a confusion matrix.

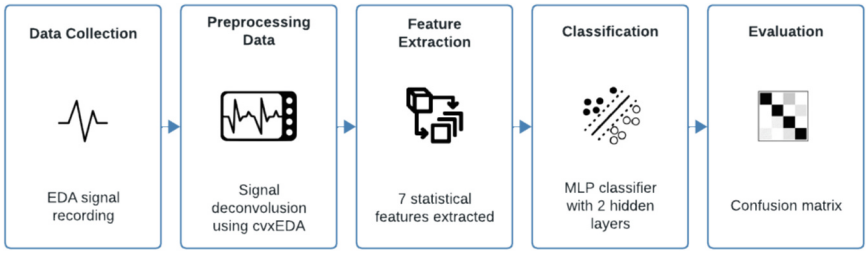


Fig. 1. Research procedures

2.1 Dataset Collection

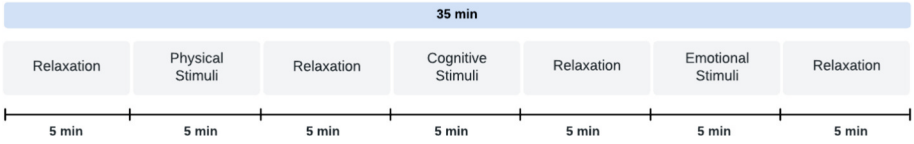


Fig. 2. Dataset recording mechanism

The data source used is a dataset from previous research [21] taken from the Public Dataset Open Access on the website (https://personal.utdallas.edu/~nourani/Bioinformatics/). In the study, data was collected in the form of SpO_2, HR, and EDA signals. Data was measured from 20 University of Texas at Dallas students consisting of 14 men and 6 women (19–33 years old) in good health (no mental disorders). In the description of the dataset, the signals were recorded while the subjects performed 7 activities consecutively. Figure 2 shows an illustration of the activities performed as stimulus according to the type of stress.

In the physical stimulus, the subject will be asked to stand for 1 min, walk on a treadmill at a speed of 1 mile/h (± 1.609 km/h) for 2 min, then walk/run on a treadmill at a speed of 3 miles/h (± 4.828 km/h) for 2 min. in the cognitive stimulus, the subject counts down the subtraction of seven in sequence, starting from the number 2485 for 3 min. Furthermore, conducting a strop test for 2 min, namely identifying the names of colors written in different colors of ink, then mentioning the color of the ink. There was

a warning bell that would ring if the subject made a mistake. In the emotional stimulus, the subject was told that a 5-min clip of a horror movie would be shown in 1 min. After 1 min of anticipation, a clip from the horror/action zombie apocalypse movie, The Horde (2009) was shown. Each stimulus will be paused for 5 min of relaxation or rest. During the 4 relaxation sessions, the subjects were asked to sit and listen to calm music commonly used in meditation. The purpose of these relaxation sessions was to establish a baseline of the physiological metrics being measured. That way it could be seen how each metric changed over the course of performing the three activities, as well as what the sensitivity and specificity of each metric was.

2.2 Pre-processing

Some of the steps performed in data pre-processing are data slicing, signal deconvolution and data segmentation. Data slicing is done to separate stress data and relaxation data, because this study only uses stress data (physical, cognitive, and emotional) without relaxation data. Next, perform signal deconvolution using the Convex Optimization (cvxEDA) method. The deconvolution is done to separate the raw Skin Conductance (SC) signal data into tonic (SCL) and phasic (SCR) components [24]. In this study, the phasic component (SCR) is used for the classification process because it describes changes in the subject's physiological response to the stimulus.

The cvxEDA model of this study uses the method proposed by [25], by decomposing the SC signal (y) into phasic (r), tonic (t), and noise components ($\in$) (Eq. 1).

$$y = r + t + \in \tag{1}$$

The single phasic response (r) is modeled using a biexponential impulse response function, called the Bateman function (Eq. 2):

$$h(\tau) = \left(e^{\frac{\tau}{\tau_0}} - e^{\frac{\tau}{\tau_1}}\right)u(\tau) \tag{2}$$

where the time constant are represented by τ_0 and τ_1 and unit step function is represented by $u(\tau)$. By utilizing autoregressive moving-average (ARMA) model of Eqs. (3) and (4) into ARMA cascade,

$$q = A^{-1}p, r = Mq \tag{3}$$

$$r = MA^{-1}p \tag{4}$$

where p represents the SMNA, MA are tridiagonal matrix. Tonic component (t) is represented as the cubic B-spline basis function (Eq. 5)

$$t = B\lambda + C \tag{5}$$

where B is a matrix for cubic B-spline basis function, λ is the vector of spline coefficients, C is a $N \times 2$ matrix with $C_{i,1} = 1$ and $C_{i,2} =$ i/N. Equations (4) and (5) might be substituted into Eq. (1) to form the observation model. Finding the maximal posteriori

spike of p and t for the recorded EDA signals is the aim. By substituting these terms into Eq. (6)

$$P\left[q, \lambda, d|y\right] \propto P\left[y|q, \lambda, d\right]P\left[q\right]P[\lambda]P[d] \tag{6}$$

and taking the algorithm of Eq. (7), the convex optimization that represents cvxEDA is:

$$\text{minimize } \frac{1}{2}\text{Mq} + \text{B}\lambda + \text{Cd} - \text{y}_2^2 + \alpha\text{Aq}_1 + \frac{\gamma}{2}\lambda_2^2 \tag{7}$$

$$\text{Subj. to } Aq(x) \geq 0 \tag{8}$$

After obtaining the phasic component data from the deconvolution method, this component will then go through a time segmentation process. This segmentation aims to multiply the data that will be used as input in the classification process.

2.3 Features Extraction

Feature extraction aims to extract parameters that represent discriminative information. This study uses 7 statistical features to characterize the results of the phasic component (SCR) that has been obtained from the results of signal deconvolution with the cvxEDA method previously described. These features include:

Mean. Mean can be defined as the average value of data that can represent the value of the data set. The equation used to calculate the mean value is found in Eq. (9) where is $\sum_{n=1}^{N} X_n$ the sum of all data and N is the sum of all data frequencies.

$$\mu = \frac{1}{N}\sum_{n=1}^{N} X_n \tag{9}$$

Standard Deviation. The standard deviation feature is the root of the sum of the squared ranges between the mean value and the highest to lowest frequency (close to zero). Standard deviation can be calculated using Eq. (10).

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N} (X_n - \mu)^2} \tag{10}$$

First Absolute Difference (FAD). This feature is an average representation for the absolute value of the raw first difference signal, where $\sum_{n=1}^{N-1} |X_{n+1} - X_n|$ is the absolute sum of all data and N is the sum of all data frequencies. This feature can be calculated by Eq. (11).

$$\delta = \frac{1}{N-1}\sum_{n=1}^{N-1} |X_{n+1} - X_n| \tag{11}$$

Normalized First Absolute Difference (NFAD). Normalization of absolute value differences can be generated through Eq. (12) below, where $\overline{x}_n$ is the normalized data.

$$\overline{\delta} = \frac{1}{N-1}\sum_{n=1}^{N-1} \left|\overline{X}_{n+1} - \overline{X}_n\right| \tag{12}$$

Rise Time. Rise time is the time a signal is measured starting from a low level (0%) to reach a high level (100%). The magnitude of the rise time value is expressed in Eq. (13), where t_2 states the time when the signal reaches the maximum level (100%) and t_1 is the time when the signal is at the minimum point (0%).

$$t_r = t_2 - t_1 \tag{13}$$

Fall Time. The opposite of rise time, fall time is the time it takes for a signal to decrease from a high value (100%) to a low value (0%). Therefore, Eq. (14) to calculate this feature is the inverse of the rise time equation.

$$t_f = t_1 - t_2 \tag{14}$$

Mean Inter-peak Interval. This feature is used to calculate the average time interval between peaks in the signal to provide information about the degree of change in the signal. The mean inter-peak interval value can be calculated using Eq. (15) below, where N is the number of peaks in the signal and Δt_i is the time interval between the ith peak and the peak of $(i - 1)$.

$$\mu_{int} = \frac{1}{\frac{1}{N}\sum_{n=1}^{N}(\Delta t_i)} \tag{15}$$

After that, the results of all the features above are normalized to have the same range for each feature. Normalization is done with Eq. (16).

$$Y = \frac{y - \min(y)}{\max(y) - \min(y)} \tag{16}$$

2.4 Classification and Evaluation

In this research, the classifier method used is Multi-Layer Perceptron (MLP). MLP basically consists of an input layer, one or more hidden layers, and an output layer [26, 27]. Classification is done for 3 classes of stress types: physical stress, cognitive stress and emotional stress. The data is first divided into training data and testing data with 80% training data and 20% test data. The result analysis stage is needed to determine the performance of the classification model with the Multi-Layer Perceptron (MLP) classifier method. In this research, a confusion matrix is used to make it easier to find the performance of the classification results. An easy-to-understand table called a confusion matrix allows us to examine the individual rows, columns, or entries to view the comprehensive classification outcomes of an algorithm on a test set [28].

There are several parameters that can be obtained with a confusion matrix, some of which are accuracy and F1-score values which contain precision and recall information. Accuracy is used to measure the model's ability to make correct predictions from all total predictions. While F1-score is used to determine the balance between the precision and sensitivity values of the classification model [29]. Precision is used to measure the model's ability to predict the correct positive class from the total positive predictions while sensitivity is used to describe how well the model correctly identifies the positive class [30].

3 Result and Discussion

The data slicing stage in EDA signal preprocessing aims to separate stress data and relaxation data. The results show that there is a change in the duration of the data length, from 2100 s to 900 s due to only stress data being used. Subsequently, the sliced signal underwent signal deconvolution to extract the SCR signal. The SCR signal only exists in a certain period and represents the subject's stress response to a given stressor [23, 31]. Figure 3 shows the results of signal deconvolution using the cvxEDA method.

Based on Fig. 3, it can be seen that the SCR signal fluctuates more sharply than the SCL signal. This occurs because the SCR signal represents changes in the subject's physiological response to the stressor, while the SCL signal represents the basal activity of the subject's physiological response [21]. Furthermore, there is an overlap between the two SCR signals between the 55th and 75th seconds, which precludes the possibility of accurately identifying it in the original SC signal. These two signals show that the cvxEDA method is able to accurately separate the SCR signal from the original skin conductivity signal data. The SCR signal obtained is then subjected to data segmentation in order to multiply the data with a segmentation duration of 60 s, 30 s and 20 s so as to produce data as much as 300, 600, and 900 data respectively.

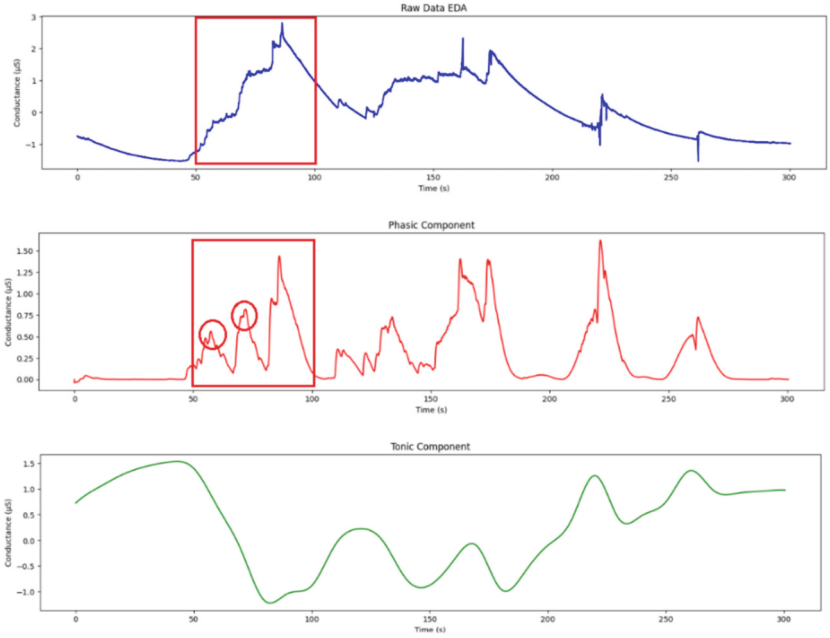


Fig. 3. The comparison of the original skin conductivity signal (upper row) with the SCR signal (middle row) and SCL signal (lower row).

Each data segment variation will have its features extracted to obtain statistical information such as mean, standard deviation, First Absolute Difference (FAD), Normalized

First Absolute Difference (NFAD), rise time, fall time, and mean inter-peak interval. The seven feature values were normalized to obtain values with a range from 0 to 1 and labeled according to the type of stress. The three normalized data variations were correlated using Pearson correlation to determine the relationship between two variables [32]. The results of statistical analysis using Pearson correlation can be seen in Fig. 4.

Figure 4 displays a diverse correlation between the extracted features and the type of stress. The highest correlation results are achieved with the 300 data segment variation, as smaller segmentation durations result in less information retrieved from the SCR signal. The data indicate a strong positive correlation between the fall time and standard deviation features, suggesting that the skin response to stressors that can increase stress plays a significant role. In contrast, the other five features showed a negative correlation indicating a decrease in EDA signal complexity at higher stress types and levels. There are 4 main features with the highest Pearson correlation values including fall time, rise time, NFAD, and mean inter-peak interval. These features represent key physiological indicators in assessing an individual's stress response. Therefore, a segmentation duration of 60 s (300 data) is able to optimize information retrieval from the SCR signal. These features represent key physiological indicators in assessing an individual's stress response. Based on the results of the correlation test, variations are made to the features used, namely the 4 features with the highest correlation and the seven features. All feature variations will be input to the classification system with 80% training data and 20% testing data.

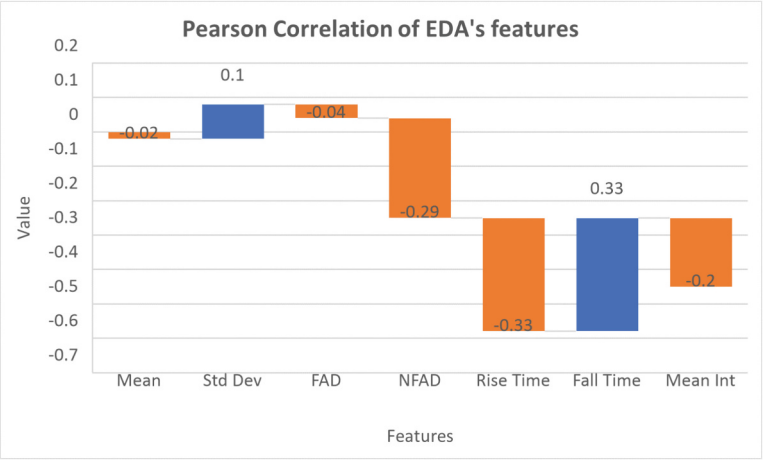


Fig. 4. Pearson correlation results on each feature

The classification system in this study uses the Multi-Layer Perceptron (MLP) algorithm with a relu activation function and a learning rate of 0.001 with 1000 epochs. In this study, the best model parameters were searched before obtaining the final model with the best performance. The results of the feature variation can be observed in Table 1. Based on these results, it can be observed that the accuracy obtained in the best combination of 4 features is 51.5%. These results still indicate poor accuracy for each type

of stress, particularly for physical stress. This may be due to the feature with the best correlation lacking important information needed for effective class separation. Variations in the data and differences in stress responses in each individual require additional features that are more informative to represent the characteristics of the data, therefore the accuracy results of 7 features provide a better accuracy value than using only the best features, which is 60%.

Table 1. Feature variation result

Feature variations	Types stress	Accuracy (%)	Average accuracy (%)
Rise time + fall time + NFAD + mean int	Physical	43	51.5
	Cognitive	52	
	Emotional	59.5	
All features	Physical	48.5	60
	Cognitive	73.5	
	Emotional	58	

Table 2. Number of hidden layer and nodes variation

Number of hidden layer	Number of nodes	Average accuracy (%)
1	20	55
	50	57.5
	75	58.33
	150	60
	200	58.66
2	(50, 25)	60.83
	(100, 50)	63.5
	(150, 75)	64.66
	(175, 75)	70
	(200, 100)	66.5
3	(100, 50, 25)	64.33
	(150, 50, 25)	63.83
	(150, 70, 35)	66
	(200, 100, 50)	67.33
	(300, 150, 75)	66.83

Moreover, the optimal parameters are evaluated based on the variations in the number of hidden layers employed and the number of nodes utilized in each hidden layer. In regard to the number of hidden layers, three variations are presented: one hidden layer, two hidden layers, and three hidden layers. Each hidden layer will have 5 variations in the number of nodes per layer. The results of the variation in the number of hidden layers and the number of nodes can be observed based on Table 2. Based on Table 2, it can be seen that the variation of 2 hidden layers with nodes (175, 75) has the best accuracy results compared to other variations. These outcomes may be attributed to the insufficient capacity of the model with a single hidden layer to effectively extract the patterns within the data, whereas those comprising three hidden layers indicate that models with excessively complex hidden layers do not consistently yield enhanced outputs. Accordingly, the classification outcomes with two hidden layers offer the most optimal results.

Once the optimal parameters for the classification system have been identified, the training and testing process is repeated five times. This is done because the MLP algorithm is a stochastic classification method where the predicted output results can provide varying values. Parameter adjustments to the classification model can play an important role in the classification results, especially in the hidden layer parameters. Adjusting the number of hidden layers as well as the number of nodes in each layer is an important factor before computation. In addition, differences in the distribution of data in each class during the testing and training process can also affect the classification results. An unbalanced class distribution can introduce bias in the model and reduce classification performance.

Table 3. Evaluation result

Stress class	Parameters	Number of repetitions (%)					Mean
		1	2	3	4	5	
Physical	Accuracy	55	60	55	65	65	60
	F1-score	61.11	68.57	62.85	65.54	68.42	65.29
Cognitive	Accuracy	85	80	85	80	80	82
	F1-score	79.11	80	83.52	78.05	80	80.13
Emotional	Accuracy	75	85	80	75	80	79
	F1-score	73.17	75.56	74.41	73.16	76.18	74.5

The next stage is the evaluation stage with test data using a confusion matrix. Based on the evaluation results described in Table 3, the average accuracy value for the three stress classes is 73.67%. The best evaluation results were obtained for the cognitive stress class, with accuracy and F1-score values of 82% and 80.13%, respectively. While the other two classes have accuracy and F1-score values of 79% and 74.5% and 60% and 65.29% for emotional stress and physical stress classes respectively.

In this study, the class with the lowest accuracy level and is below the good accuracy standard is physical stress, with an average accuracy of 60%. The low classification

results may be due to differences in the physiological response of the subject to the given stimulus or stressor. In the case of physical stress, the stimulus given is walking and running on a treadmill, which may have different effects on each subject. It is known that the subjects consisted of 14 males and 6 females with healthy conditions where differences in biological and physiological characteristics may affect their response to physical stimuli.

In addition to differences in biological and physiological characteristics, the subject's ability to respond to a stimulus may decrease if it has been done repeatedly, this can be referred to as habituation [33]. Habituation describes a progressive decrease in the body's physiological response when a person is repeatedly exposed to the same stimulus, so that the stimulus is no longer considered a significant threat or stressor by the body [34]. When the subject runs on a treadmill for the first time, the EDA response will increase in response to the unusual physical activity. However, if the subject is used to doing the stimulus, the body has adapted and gives a lower response. This habituation can affect the results of stress classification as the de-creasing EDA response has the potential to blur the distinction between stressed and unstressed conditions.

Based on Table 3, cognitive stress has the highest accuracy value reaching 82% compared to the other two types of stress. While in [21] research using the same dataset, the physical stress class ranked the highest. This is due to additional information from several sensors such as HR and SpO_2 that can evaluate the body's physiological response to the stress given so that there are differences in results with previous research. Therefore, it can be concluded that in this study the use of EDA sensors is more suitable for classifying cognitive stress. Good results in classifying the type of stress in the future can provide stressful situations as early as possible so that they can provide treatment according to the type of stress they have.

4 Conclusion

This research aims and focuses on using a combination of the cvxEDA method with MLP so that it is expected to provide better deconvolution results in order to get more optimal classification results. The combination of deconvolution methods using cvxEDA and classification with MLP (Multilayer perceptron) provides a better accuracy value in the classification of cognitive stress type. The model achieved an accuracy of 82% and an F1-score of 80.13% for the cognitive stress type. These results indicate that MLP models applied to EDA signals have the potential to detect stress, especially cognitive stress and prevent other health problems. The good results of classifying types of stress in the future can provide stressful situations as early as possible so that they can provide treatment according to the type of stress.

References

1. WHO. Mental health. https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response. Accessed 16 July 2024
2. CDC. About mental health. https://www.cdc.gov/mentalhealth/learn/index.htm. Accessed 15 July 2024

3. WHO. Mental disorders. https://www.who.int/news-room/fact-sheets/detail/mental-disorders/. Accessed 16 July 2024
4. Davis MT, Holmes SE, Pietrzak RH, Esterlis I (2017) Neurobiology of chronic stress-related psychiatric disorders: evidence from molecular imaging studies. Chron Stress 1. https://doi.org/10.1177/2470547017710916
5. Sanford LD, Suchecki D, Meerlo P (2015) Stress, arousal, and sleep. Natl Libr Med 379–410. https://doi.org/10.1007/7854_2014_314
6. Sheng JA et al (2021) The hypothalamic-pituitary-adrenal axis: development, programming actions of hormones, and maternal-fetal interactions. Front Behav Neurosci 14:1–21. https://doi.org/10.3389/fnbeh.2020.601939
7. Li Y et al (2019) Differences of physical vs. psychological stress: evidences from glucocorticoid receptor expression, hippocampal subfields injury, and behavioral abnormalities. Brain Imaging Behav 13:1780–1788. https://doi.org/10.1007/s11682-018-9956-3
8. Stults-Kolehmainen MA, Sinha R (2014) The effects of stress on physical activity and exercise. https://doi.org/10.1007/s40279-013-0090-5
9. Ingegnoli F, Castelli R, Gualtierotti R (2013) Rheumatoid factors: clinical applications, vol 35, pp 727–734
10. Folkman S, Lazarus RS (2013) Stress and coping theory applied to the investigation of mass industrial psychogenic illness. New York
11. APA. Stress effects on the body. https://www.apa.org/topics/stress/body. Accessed 20 July 2024
12. Pop-Jordanova N, Pop-Jordanov J (2020) Electrodermal activity and stress assessment. Prilozi 41:5–15. https://doi.org/10.2478/prilozi-2020-0028
13. Braithwaite JJJ et al (2013) A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. In: CTIT technical reports series, pp 1–42
14. Cho D et al (2017) Detection of stress levels from biosignals measured in virtual reality environments using a kernel-based extreme learning machine. Sensors (Switzerland) 17. https://doi.org/10.3390/s17102435
15. Boucsein W (2012) Electrodermal activity. Springer, New York
16. Posada-Quintero HF, Chon KH (2020) Innovations in electrodermal activity data collection and signal processing: a systematic review. Sensors (Switzerland) 20. https://doi.org/10.3390/s20020479
17. Nardelli M, Greco A, Sebastiani L, Scilingo EP (2022) ComEDA: a new tool for stress assessment based on electrodermal activity. Comput Biol Med 150. https://doi.org/10.1016/j.compbiomed.2022.106144
18. Zontone P, Affanni A, Bernardini R, Piras A, Rinaldo R (2019) Stress detection through electrodermal activity (EDA) and electrocardiogram (ECG) analysis in car drivers. In: European signal processing conference. https://doi.org/10.23919/EUSIPCO.2019.8902631
19. Mozafari M, Khazaei E, Bagheri S, Hasanzadeh N, TaghiBeyglou B, Ghassemi F (2019) Mental stress level detection using physiological signals and evolutionary algorithm. In: 2019 26th national and 4th international Iranian conference on biomedical engineering (ICBME). https://doi.org/10.1109/ICBME49163.2019.9030424
20. Ileri R, Latifoglu F (2020) Analysis of the electrodermal activity signals for different stressors using empirical mode decomposition. Acad Platform J Eng Sci 407–414. https://doi.org/10.21541/apjes.601235
21. Birjandtalab J, Cogan D, Pouyan MB, Nourani M (2016) A non-EEG biosignals dataset for assessment and visualization of neurological status. In: IEEE workshop on signal processing systems, SiPS: design and implementation, pp 110–114. https://doi.org/10.1109/SiPS.2016.27

22. Greco A et al (2023) Acute stress state classification based on electrodermal activity modeling. IEEE Trans Affect Comput 14. https://doi.org/10.1109/TAFFC.2021.3055294
23. Rahma O et al (2022) Electrodermal activity for measuring cognitive and emotional stress level. J Med Signals Sens 12. https://doi.org/10.4103/jmss.JMSS_78_20
24. Greco A, Valenza G, Lanata A, Scilingo EP, Citi L (2016) CvxEDA: a convex optimization approach to electrodermal activity processing. IEEE Trans Biomed Eng 63:797–804. https://doi.org/10.1109/TBME.2015.2474131
25. Greco A, Valenza G, Scilingo EP (2016) Advances in electrodermal activity processing with applications for mental health: from heuristic methods to convex optimization. https://doi.org/10.1007/978-3-319-46705-4
26. Lattanzi E, Donati M, Freschi V (2022) Exploring artificial neural networks efficiency in tiny wearable devices for human activity recognition. Sensors 22. https://doi.org/10.3390/s22072637
27. Dastres R, Soori M, Neural A, Systems N, Journal I (2021) Artificial neural network systems to cite this version: HAL IAd: hal-03349542. Int J Imaging Robot (IJIR) 21:13–25
28. Singh P, Singh N, Singh KK, Singh A (2021) Diagnosing of disease using machine learning. In: Machine learning and the internet of medical things in healthcare, pp 89–111. https://doi.org/10.1016/B978-0-12-821229-5.00003-3
29. Hand DJ, Christen P, Kirielle N (2021) F*: an interpretable transformation of the F-measure. Mach Learn 110:451–456. https://doi.org/10.1007/s10994-021-05964-1
30. Hicks SA et al (2022) On evaluation metrics for medical applications of artificial intelligence. Sci Rep 12:1–9. https://doi.org/10.1038/s41598-022-09954-8
31. Hernando-Gallego F, Luengo D, Artes-Rodriguez A (2018) Feature extraction of galvanic skin responses by nonnegative sparse deconvolution. IEEE J Biomed Health Inform 22:1385–1394. https://doi.org/10.1109/JBHI.2017.2780252
32. Hidayanti AA, Dewi EN, Mandalika (2023) Pearson correlation analysis of production costs on the land area of salt farmers in bolo sub-district, Bima district. J Inov Pend Sains 4:5–9
33. Sarchiapone M et al (2018) The association between electrodermal activity (EDA), depression and suicidal behaviour: a systematic review and narrative synthesis. BMC Psychiatry 18:1–27. https://doi.org/10.1186/s12888-017-1551-4
34. Dissegna A, Turatto M, Chiandetti C (2021) Context-specific habituation: a review, pp 1–11

Efficient Gradient Boosting Integrated with Advanced Text Vectorization for Pulmonary Disease Classification

Phuoc-Hai Huynh[1,2(✉)]

[1] Faculty of Information Technology, An Giang University, An Giang, Vietnam
hphai@agu.edu.vn
[2] Vietnam National University, Ho Chi Minh City, Vietnam

Abstract. We presents a novel methodology for classifying pulmonary diseases using Vietnamese clinical text data, tackling issues of data complexity and inconsistency in electronic medical records, along with the inadequate applicability of English-centric models. Clinical data from An Giang Regional General Hospital were transformed into feature vectors using Embedding-based and Large Language Models (LLMs). The experimental outcomes demonstrate that the XGBoost model, when combined with Term Frequency–Inverse Document Frequency (TF-IDF) and Bag of Words (BoW) representations, achieved the highest classification accuracies of 87.66% and 87.63%, respectively, thereby outperforming the other evaluated models. Significant contributions encompass the development of a standardized Vietnamese clinical dataset for 12 pulmonary diseases and an illustration of the potential for integration into hospital information systems to improve diagnostic precision and healthcare efficiency.

Keywords: pulmonary diseases · classification models · electronic medical records · XGBoost · BoW · TF-IDF · machine learning

1 Introduction

Pulmonary disease represents a major contributor to worldwide mortality and morbidity, which includes lung cancer, asthma, pneumonia, tuberculosis, and chronic obstructive pulmonary disease [1]. In the Global Asthma Report [2], pulmonary diseases affect approximately 262 million people globally and result in over 1,000 deaths daily, many of which are preventable five, pneumonia accounts for 14% of all deaths, and tuberculosis aimed 1.6 million lives in 2021.

Healthcare is one of the many industries in which information technology has led to significant advances. In Vietnam, information management systems for hospitals are used widely. Exam data, treatment records, and electronic medical records (EMRs) are all stored with these systems. Clinical data is recorded by EMRs and stored in databases. Clinical signs include symptoms that patients report as well as signs that are found through exams. Clinical exams evaluate basic health status, assess risk factors for

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 160–171, 2026.
https://doi.org/10.1007/978-981-95-9233-3_13

disease, and determine disease causes. The amount of data in EMRs is growing quickly. In international medicine, research and analysis of EMRs data are popular trends. EMRs are huge data stores. It makes it possible to access recorded knowledge and to explore and update new medical knowledge [3]. The mining and analysis of EMRs data aids machine learning models. These models can improve healthcare, automate diagnosis, and optimize the treatment procedure.

This study presents a highly efficient approach for classifying pulmonary diseases. This method utilizes Vietnamese clinical text data obtained from electronic medical records. At present, the diagnosis and classification of pulmonary diseases face a multitude of challenges [4, 5]. These obstacles appear from the presence of inconsistent and complex clinical data that is hard to analyze. Previous studies have found challenges associated with literature in the field of medicine. The problems addressed in this study are characterized by a lack of precision and inconsistent outcomes in classification [4, 5]. Traditional methods often require the involvement of medical professionals in analyzing the data. It also leads to an increase in both the time it takes to diagnose a condition and the associated costs. Therefore, conducting research and developing novel techniques for identifying pulmonary diseases from Vietnamese clinical text data is of significant importance.

The contributions of this research are threefold: First, we collected and standardized a clinical examination dataset from An Giang Regional General Hospital to ensure consistency and usability. Second, we applied various vectorization techniques, including embedding methods and large language models, to convert clinical text into feature vectors and compared several classification models. Finally, We found that the XGBoost model worked best for pulmonary disease. It gave the highest accuracy when used with Bag of Words (BoW) and Term Frequency–Inverse Document Frequency (TF-IDF). This model performed well because it can handle complex features.

The paper is structured into several sections. Section 2 provides an overview of previous studies. Sections 3 and 4 describe the proposed methodology and report the experimental outcomes. Sections 5 and 6 offer a discussion of the findings and outline directions for future research.

2 Literature Review

Machine learning has shown significant potential for improving disease classification and offering enhanced diagnostic and treatment support [6]. In the context of pulmonary diseases, numerous studies have applied machine learning techniques to classify medical data, particularly text data, due to its ability to extract meaningful patterns. However, existing approaches often rely on English-language datasets, limiting their applicability in Vietnamese healthcare. This review examines recent research on disease classification using supervised learning techniques and language models, highlighting their strengths and limitations.

Supervised learning techniques have proved effective in classifying diseases based on text data [6]. The diagnostic model is studied for peripheral neuropathy using fuzzy logic. The experiment, based on 24 million characteristic symptoms and paraclinical examination information, achieved 93.27% accuracy compared to expert diagnoses [7]. In a study [8], the authors used machine learning algorithms to classify clinical text reports into four diseases. The experimental outcomes indicate that both the linear model and the probabilistic classifier, representing statistical learning approaches, outperformed the remaining models with an accuracy of 96.2%.

Related studies have shown the benefits of using vectorization models and LLMs. On the one hand, these models can convert electronic medical record data into numerical features. The study [9] conducted a study using BoW and TF-IDF. They used these methods to classify medical texts. Their research showed that TF-IDF provided higher accuracy. This higher accuracy was observed in identifying specific medical conditions. In addition, LLMs are applied to process and classify electronic medical records. BERT is used in the study as a method for extracting relevant information from clinical notes [10].

Although promising, most algorithms have been developed using data from other countries, primarily in English. This highlights a gap in the research for Vietnamese datasets, where the specific characteristics of the Vietnamese language and medical terminology create unique challenges. This study addresses this gap using on Vietnamese clinical data and developing methods tailored to its linguistic and contextual nuances.

3 Study Design and Methods

3.1 Data Processing

In order to build a pulmonary disease classification dataset, we collected paraclinical examination information from the Hospital Information Management (HIS) system of An Giang Provincial General Hospital. The novel dataset has 17,879 text files. It is labeled with 12 common pulmonary diseases, including pneumonia, COVID-19, tuberculosis, COPD, asthma, pulmonary edema, respiratory failure, effusion, emphysema, pulmonary tumor, atelectasis, and normal. We have chosen suitable characteristics to be used in the following machine learning models, since our objective is to classify pulmonary diseases by utilizing clinical examination text information. The dataset consists of 17,879 records after data cleaning and appropriate attribute selection. Each record contains seven attributes: summary of para-clinical results, general physical examination, breathing system, cardiovascular system, personal medical history, and medical record summary. In Table 1, the clinical dataset is summarized, detailing the particular attributes essential for the classification of pulmonary diseases.

Table 1. Illustrative example of clinical text extract from EMR

Clinical symptoms	General physical examination	Respiratory examination	Cardiovascular examination	Mental status examination	Laboratory and imaging findings
Symptoms onset 2 days ago, chest pain, back pain radiating to the thoracic area (D7), increased intensity with movement, intermittent	Patient conscious, pale skin, SpO_2: 94%, BP: 120/80 mmHg	Breathing normal, crackles noted	Pulse regular, tachycardia	Blood pressure normal, no fever. Diagnosis: chest pain due to musculoskeletal causes	Chest X-ray findings: T4 vertebral degeneration
Patient 15 years old, fatigue and fever for 1 day	Conscious, tired, fever 38 °C. Reddened conjunctiva	Respiratory function reduced	Heart sounds normal	Blood pressure elevated. Diagnosis: type 2 diabetes mellitus with hypertension	Blood glucose elevated: 11.05. WBC: 15.1, MONO%: 81.09, NEU%: 14.20. X-ray: mild pleural effusion
Complaints of diffuse chest pain, unrelated to exertion, not aggravated by rest	Conscious, tired, BP 140/80 mmHg, pain score 6/10	Breathing normal	Heart rate regular	No abnormal findings in mental status	No abnormalities found in imaging and tests

Table 2 provides the distribution of cases across pulmonary disease categories, with the dataset divided into two subsets: 80% used to train and 20% for testing.

Table 2 and Fig. 1 show that pneumonia is the most common pulmonary disease. COVID-19 also has a significant number of cases. Diseases such as respiratory failure and pulmonary tuberculosis have a moderate prevalence, while other diseases such as atelectasis, pulmonary cancer, and pulmonary edema have lower numbers of cases. This chart helps us better understand the distribution and prevalence of different types of pulmonary diseases.

Table 2. Distribution of training and testing datasets for various pulmonary diseases

Diseases	Training datasets	Testing dataset
Pneumonia	6472	1618
Normal	2469	618
COVID-19	2000	501
Respiratory failure	1105	277
Pulmonary tuberculosis	657	165
COPD	490	123
Pleural effusion	452	114
Pneumothorax	213	54
Asthma	153	39
Pulmonary edema	149	38
Pulmonary cancer	128	33
Atelectasis	8	3

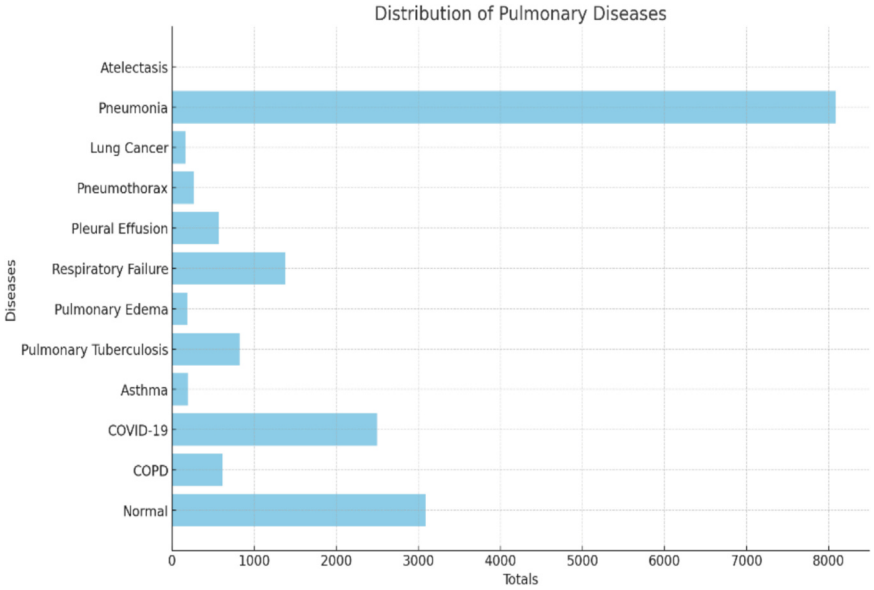


Fig. 1. Distribution of pulmonary diseases

3.2 Embedding-Based and LLM Approaches for Pulmonary Disease Classification

In the process of building a pulmonary disease classification model, the process is divided into three stages, as shown in Fig. 2. First, the text-based medical examination data is

preprocessed to remove noise. Next, we use vectorization methods using two groups of techniques. The group of methods of embedding models, including BoW, TF-IDF, Word2-Vec [10] and GloVe [11] and LLMs, including BERT [12], MiniLM-L6-L12 [13], ViT5 [14], and DistilBERT [15]. As a result of this step, the medical examination data is converted into vectors and labeled with the diseases in the hospital discharge records. Finally, nine classifiers are used to classify and evaluate.

Data preprocessing comprises several steps, including the removal of extraneous characters, punctuation, numbers, and special symbols; the elimination of stop words; the conversion of text to lowercase; and the application of tokenization. Next, we vectorize the dataset using embedding-based and LLMs approaches to transform the textual data into numerical vectors.

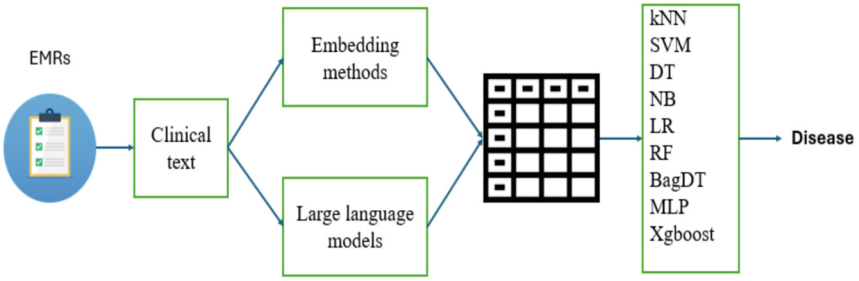


Fig. 2. Workflow for classifying lung diseases using clinical text data

Embedding methods are effective approaches to converting raw clinical text data into number vectors, which makes it easier to use them in machine learning models for classifying tasks. These techniques convert text data with numerous dimensions into continuous vector spaces with smaller dimensions. This change helps you record the semantic meanings and relationships among words and phrases in the text data. The techniques BoW, TF-IDF, Word2Vec, GloVe, and BERT are all effective methods for representing medical text data as vectors used for classification and prediction in machine learning models.

Additionally, in this study, we also use LLMs, including MiniLM-L6, MiniLM-L12, ViT5, BertBase, and DistilBERT, for vectorization. Over the past few years, LLMs have become indispensable resources in advancing natural language processing and artificial intelligence. The pre-trained models is used to vectorize clinical text data, without training additional data for the model.

For the classification task, nine models were utilized: *k*NN, decision trees (J48), SVM, random forest (RF), bagging of decision trees (Bag-DT), multi-layer perceptron (MLP), logistic regression (LR), and extreme gradient boosting (XGBoost).

4 Evaluation

4.1 Experiments Setup

A total of nine models were implemented to perform classification, comprising *k*NN, SVM, DT, RF, Bag-DT, MLP, LR, and XGBoost. Scikit-learn framework [27] is used in experiment. For LLMs, we use the hugging face library [33]. The experiments were performed on a Ubuntu with GPU 1080-Ti. Model performance was quantitatively evaluated using overall classification accuracy as the principal metric. Furthermore, the classification outcomes of the learning algorithms were statistically examined through Student's *t*-test.

4.2 Classification Results

As shown in Table 3, classification performance is reported using accuracy. The classification outcomes are further visualized through the plots in Figs. 3 and 4. Columns 2–10 detail the results obtained from the remaining algorithms. In addition, Table 5 provides the paired statistical test results.

Table 3. Accuracy (%) of 9 classifiers with 5 embedding methods.

	kNN	SVM	DT	NB	RF	BagDT	MLP	LR	XGBoost
BoW	76.06	83.56	78.63	46.26	83.61	85.53	84.51	83.92	**87.63**
Glove	70.54	78.39	64.87	37.42	77.35	77.44	78.16	69	80.67
TFIDF	68.12	84.38	79.38	46.28	82.98	85.42	82.93	81.53	**87.66**
Word2Vector	73.37	81.32	69.36	42.62	79.76	80.06	81.64	78.16	82.77
BERT	71.33	80.04	61.65	23.06	76.01	75.79	80.78	79.97	80.26

Table 4 summarizes the classification accuracy of nine models across five distinct embedding techniques. When applying TF-IDF and BoW representations, XGBoost consistently achieves the highest classification accuracy, with values of 87.66% and 87.63%, respectively. The BoW and TF-IDF embedding methods yield the best results for most classifiers, while GloVe and BERT show lower performance. In summary, choosing a suitable classifier and embedding method is critical; the best classifier for classifying Vietnamese clinical text data is XGBoost, while the best embedding methods are BoW and TF-IDF.

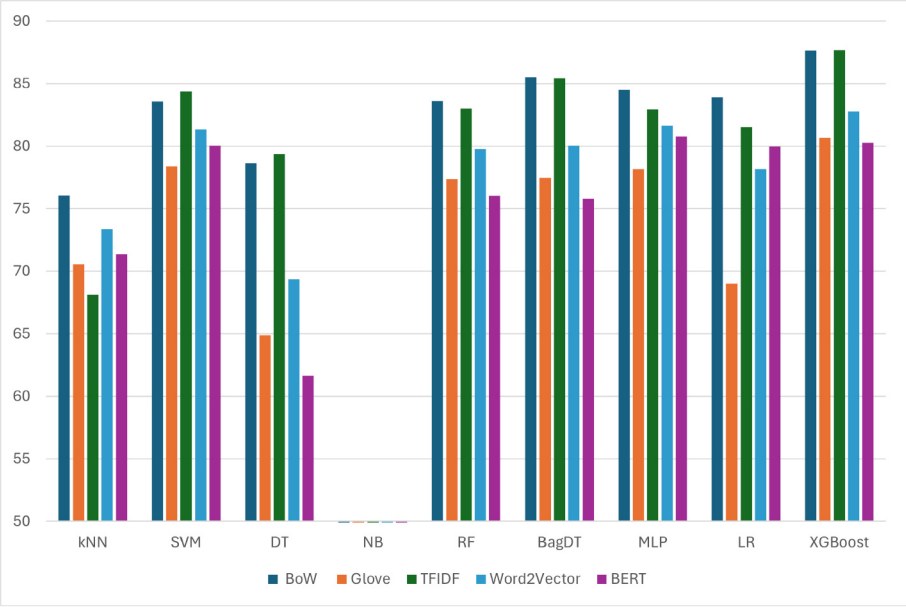


Fig. 3. Performance comparison of classifiers using five embedding methods.

Table 4. Accuracy (%) of 9 classifiers with 5 large language models (LLMs)

	kNN	SVM	DT	NB	RF	BagDT	MLP	LR	XGBoost
MiniLM-L6	65.39	71.67	60.73	36.9	72.85	73.1	73.1	65.7	75.27
MiniLM-L12	63.03	69.03	59.87	42.05	71.44	70.92	70.4	64.87	74.23
ViT5	63.12	75.49	59.64	25.37	72.51	72.82	76.01	75.27	75.72
BertBase	62.02	72.56	58.78	28.28	71.61	71.89	70.51	71.12	73.97
DistilBERT	64.35	74.93	59.81	27.82	72.5	72.26	69.95	69.74	74.87

Table 4 reports the evaluation of nine classifiers over five LLMs. Notably, XGBoost consistently outperformed the others, with its best accuracy (75.72%) obtained under the ViT5 setting. While MLP also achieves a high accuracy of 76.01%, this model performs well overall. Although MiniLM-L6 and MiniLM-L12's accuracies are marginally lower than ViT5's, they nevertheless exhibit strong performance, with XGBoost achieving 75.27% and 74.23%, respectively. In conclusion, it is critical to select the right LLM and classifier; for the purpose of classifying Vietnamese clinical text data, the combination of ViT5 and XGBoost works best.

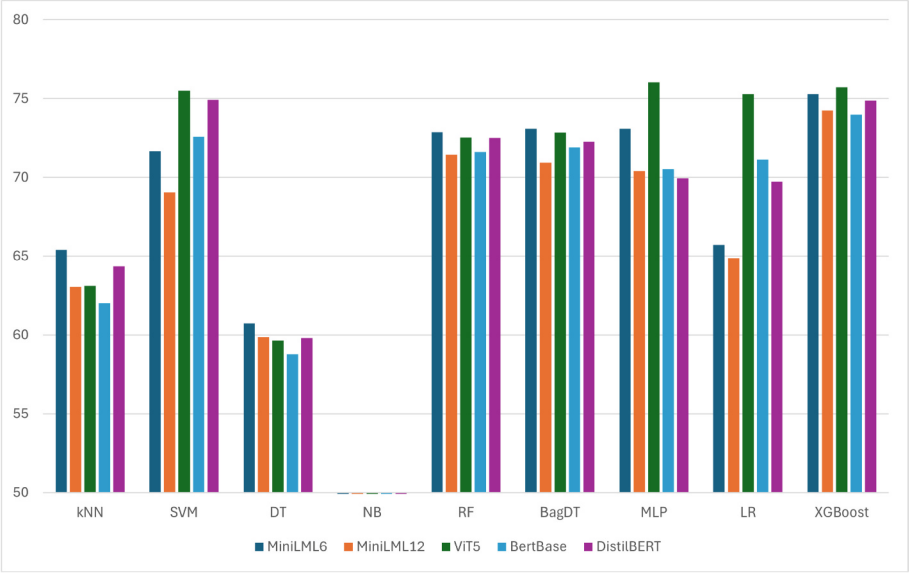


Fig. 4. Comparison of overall accuracy of classifiers on LLMs

Table 5. T-test results comparing XGBoost with other classifiers.

Comparison	p-value
XGBoost and kNN	0.00000073
XGBoost and SVM	0.00439575
XGBoost and DT	0.00000020
XGBoost and NB	0.00000001
XGBoost and RF	0.00000059
XGBoost and BagDT	0.00000089
XGBoost and MLP	0.00247890
XGBoost and LR	0.00168687

Table 5 shows the t-test p-values between the performance of XGBoost and other classifiers on all embedding methods as well as LLM models. The p-values reported in Table 5 indicate statistically significant performance differences between XGBoost and the other classifiers across all embedding methods and LLMs. All p-values are less than 0.05. This indicates that XGBoost performs better compared to the other classifiers.

Comparing the accuracy of XGBoost when classifying with BoW and TF-IDF versus large language models (LLMs) shows that XGBoost achieves the highest accuracy with BoW (87.63%) and TF-IDF (87.66%), significantly outperforming all LLMs. Among the LLMs, ViT5 achieves the best accuracy with 75.72%, followed by MiniLM-L6 (75.27%)

and DistilBERT (74.87%). The performance of XGBoost with LLMs is significantly lower compared to BoW and TF-IDF, with a difference of approximately 12% between TF-IDF and ViT5. In summary, XGBoost with BoW and TF-IDF achieves the highest performance, indicating that traditional embedding methods are more effective than LLMs in classifying Vietnamese clinical text data.

5 Discussion

This section provides a detailed discussion of the results, compares them with findings from other studies, identifies the limitations of the present work, and outlines its implications.

Based experiment results, XGBoost model with the BoW and TF-IDF embedding methods outperforms other classifiers and embedding methods in classifying pulmonary diseases on Vietnamese clinical text data. In particular, XGBoost achieved 87.66% accuracy with TF-IDF and 87.63% with BoW, demonstrating its ability to optimize learning and handle complex features. These findings suggest that BoW and TF-IDF, which convert text data into numerical vectors while retaining semantic meaning, are better for this dataset than more advanced large language models. SVM and Bagging of Decision Trees also performed well, but not as well as XGBoost.

In Vietnamese clinical text data, the proposed method outperforms others. Previous research used open datasets mostly in English, which may not capture Vietnamese medical terminology. In Vietnamese clinical text data, the proposed method outperforms others. Previous research used open datasets mostly in English, which may not capture Vietnamese medical terminology. Similar vectorization techniques, like [16, 17], showed that TF-IDF improves medical text classification accuracy, supporting our findings. LLMs have been used in studies to classify clinical text [18]. However, in order to make them more accurate, they need to be trained more or have their structure changed to fit the data.

This study's limitation is that it depends on clinical text data quality and structure. EMR noise, inconsistencies, and incomplete records may affect classification model accuracy. Traditional embedded methods like BoW and TF-IDF performed well, but advanced LLMs may better capture medical text context and semantics. Another limitation of this study is the lack of investigation into whether fine-tuning LLMs on the Vietnamese clinical dataset could enhance performance. The results are also limited to a single hospital, which may limit their applicability to other healthcare settings in Vietnam or elsewhere.

However, there are some limitations. However, our proposal has major healthcare implications in Vietnam. This study shows that machine learning models, particularly XGBoost combined with traditional embedding methods, can classify pulmonary diseases from clinical text data, improving diagnostic accuracy and efficiency in Vietnamese hospitals. Integrating the method into HIS helps doctors make faster, more accurate diagnoses, improving patient care. This approach can be applied to other clinical data and diseases, making it a versatile tool for medical research.

6 Conclusion and Future Works

In this work, we develop a practical approach to classify pulmonary diseases utilizing Vietnamese clinical narratives obtained from EMRs. Results show that XGBoost, combined with traditional vectorization methods like BoW and TF-IDF, achieves the highest classification accuracy, reaching 87.66% and 87.63%, respectively. These findings confirm XGBoost's effectiveness with these vectorization techniques for complex clinical data. Future research will fine-tuning LLMs for Vietnamese clinical text, expanding the dataset to include records from multiple hospitals, and exploring applications to other clinical data types and diseases. Moreover, integrating the system into hospital information systems and validating it through real-world trials will be essential for practical use.

Acknowledgment. This research is funded by Vietnam National University Ho Chi Minh City (VNU-HCM) under grant number C2024-16-02.

References

1. Wang Z et al (2025) Global, regional, and national burden of chronic obstructive pulmonary disease and its attributable risk factors from 1990 to 2021: an analysis for the Global Burden of Disease Study 2021. Respir Res 26:2
2. Yuan L et al (2025) Global, regional, national burden of asthma from 1990 to 2021, with projections of incidence to 2050: a systematic analysis of the global burden of disease study 2021. EClinicalMedicine 80
3. Ye J, Woods D, Jordan N, Starren J (2024) The role of artificial intelligence for the application of integrating electronic health records and patient-generated data in clinical decision support. AMIA Summits Transl Sci Proc 2024:459
4. Zhang S, Metaxas D (2024) On the challenges and perspectives of foundation models for medical image analysis. Med Image Anal 91:102996
5. Gao S et al (2021) Limitations of transformers on clinical text classification. IEEE J Biomed Health Inform 25:3596–3607
6. Mekov E, Miravitlles M, Petkov R (2020) Artificial intelligence and machine learning in respiratory medicine. Expert Rev Respir Med 14:559–564
7. Kunhimangalam R, Ovallath S, Joseph PK (2014) A clinical decision support system with an integrated EMR for diagnosis of peripheral neuropathy. J Med Syst 38:38
8. Khanday AMUD (2020) Machine learning based approaches for detecting COVID-19 using clinical text data. Int J Inf Technol 12:731–739
9. Le T-D, Noumeir R, Rambaud J, Sans G, Jouvet P (2021) Machine learning based on natural language processing to detect cardiac failure in clinical narratives. arXiv E-prints. arXiv-2104
10. Zhou W, Miller TA (2023) Improving the transferability of clinical note section classification models with BERT and large language model ensembles. In: Proceedings of the conference. Meeting. Association for Computational Linguistics, p 125
11. Pennington J, Socher R, Manning CD (2014) Glove: global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1532–1543

12. Devlin J, Chang M-W, Lee K, Toutanova K (2019) BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers), pp 4171–4186
13. Wang W, Wei F, Dong L, Bao H, Yang N, Zhou M (2020) MiniLM: deep self-attention distillation for task-agnostic compression of pre-trained transformers. Adv Neural Inf Process Syst 33:5776–5788
14. Phan L, Tran H, Nguyen H, Trinh TH (2022) ViT5: pretrained text-to-text transformer for Vietnamese language generation. arXiv prepr. arXiv:220506457
15. Sanh V (2019) DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. In: Proceedings of thirty-third conference on neural information processing systems (NIPS2019)
16. Sun P, Wang L, Xia Q (2017) The keyword extraction of Chinese medical web page based on WF-TF-IDF algorithm. In: 2017 international conference on cyber-enabled distributed computing and knowledge discovery (CyberC). IEEE, pp 193–198
17. Park H, Lee M, Hwang S, Oh S (2016) TF-IDF based association rule analysis system for medical data. KIPS Trans Softw Data Eng 5:145–154
18. Pressman SM, Borna S, Gomez-Cabello CA, Haider SA, Haider CR, Forte AJ (2024) Clinical and surgical applications of large language models: a systematic review. J Clin Med 13:3041

Sampling Rate Exploration to Obtain the Most Optimal Signal Resolution on 12-Lead ECG

Priyambada Cahya Nugraha, Anita Miftahul Maghfiroh(✉),
I Dewa Putu Arvian Danuartha, Muhammad Abyan Nafianto,
Salma Syahidah Nurromadhona, Tetrik Fa'altin, and I Dewa Gede Hari Wisana

Department of Electromedical Engineering, Health Polytechnic Ministry of Health Surabaya, Surabaya, Indonesia
anitamiftah@poltekkesdepkes-sby.ac.id

Abstract. Electrocardiogram (ECG) 12-lead is a non-invasive method used to monitor the electrical activity of the heart. Errors in heart recording lead to misdiagnoses in patients. One effort to improve the quality of ECG recording is by determining the data length recording time. This study aims to find the best sampling rate used by the STM32F7 in recording 12-lead ECG signals through precordial and limb lead placement. The research methodology involves configuring the ADS1293 module to record ECG signals from electrodes at precordial lead placement. ECG signals are recorded at varying sampling rates, namely at 200Hz to 1000Hz. After being transmitted, the ECG signal is analyzed using the signal-noise ratio (SNR) method to determine the ratio of noise and the original signal. The analysis results show that the higher sampling rate on the STM32F7 can provide better ECG signal details and increase accuracy in detecting QRS complex and ST segment changes. However, it is important to remember that using higher sampling rates also requires more resources and affects data processing speed. The best sampling rate obtained was 400 Hz with a delay setting of 2s, resulting in an SNR value of 12.8386 dB. This research provides a better understanding of the effect of sampling rate on the ADS1293 module with processing using the STM32F7 in recording 12-lead ECG signals through precordial and limb lead placement. This information may aid the development of medical devices that utilize the STM32F7 to incorporate cardiovascular health, taking into account the trade-off between recording quality and resource usage.

Keywords: ECG 12-lead · STM32F7 · ADS1293 · Sampling rate

1 Introduction

An ECG is a diagnostic instrument that used to detect electrical activity in the heart through a diagram that records changes in the flow of electricity in the heart over time. In the ECG signal waveform, there are wave patterns known as PQRS and T waves [1]. The main purposes of ECG include identification of heart rhythm abnormalities, detection of problems in the heart muscle, assessment of heart performance, estimation of heart enlargement or hypertrophy in the atria and ventricles, and analysis of the effects of

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 172–183, 2026.
https://doi.org/10.1007/978-981-95-9233-3_14

cardiac drugs [2]. ECG signal recording involves the use of electrodes in two placements. Electrodes placed on a patient's wrists and ankles to measure the electrical activity of the heart from the perspective of the limbs are known as limb leads. Meanwhile, precordial leads are electrodes placed on the chest to measure the electrical activity of the heart from a thoracic perspective [3]. In the context of using the STM32F7 microcontroller, this background discusses the importance of analyzing the sampling rate on a 12-lead ECG with precordial and limb leads because the sampling rate can affect the resolution of the 12-lead ECG signal [4, 5]. The sampling rate refers to the sampling frequency of the ECG signal. An appropriate sampling rate is an important factor in maintaining signal quality and the accuracy of ECG interpretation [6]. The use of the STM32F7 microcontroller in ECG signal recording provides advantages in terms of real-time signal processing capability and device design flexibility [7]. Careful analysis of the sampling rate on a 12-lead ECG with precordial and limb leads using the STM32F7 helps understand its effect on signal quality and interpretation accuracy [8].

Factors such as sampling rate that is too low can result in less accurate information and sacrifice temporal resolution, however, such advantages do not come without compromise [9]. Using more bits also requires a longer conversion time in converting analog signals to digital, potentially limiting the achievable sampling rate [10]. On the other hand, the use of fewer bits allows for faster conversion and can increase the sampling rate, but at the cost of compromising the resolution of the measurement results [11]. Meanwhile, a sampling rate that is too high can result in excessive computational load and increased power consumption. In capturing ECG data with precordial and limb leads using the STM32F7, an analysis is required that considers technical factors such as the microcontroller's memory capacity, processing speed, and the level of noise that may occur [12]. With a good understanding of the effect of sampling rate on ECG signals, device developers and researchers can optimize sampling rate settings to meet specific needs and ensure optimal signal quality.

Through comprehensive research and analysis, information on STM32F7 sampling rate analysis on 12-lead ECG with precordial and limb leads can help improve accurate clinical interpretation and improve the design of ECG devices integrated with STM32F7 microcontrollers [13–15]. This will contribute to the development of ECG devices that are more sophisticated, efficient, and reliable in their use in daily medical practice [16, 17]. From several existing journals, some have developed research on 12-lead ECG, with analysis using a 12-channel ECG machine equipped with an ATmega microcontroller, 64x12 LCD, and can be connected to other devices. However, the ECG is not equipped with 1mV calibration, paper speed setting, sensitivity, and heart rate data retrieval is not real-time [18].

Another research produced an innovation in the form of a 12-lead Electrocardiograph specifically designed to remotely monitor heart health and has a weakness in terms of data transmission, where data loss occurs when using a slow internet connection [18–20]. The next research discussed the Computer-Based 12 Lead ECG Data Acquisition Instrumentation System. Furthermore, the next research discussed the Analysis of Sampling Frequency and Resolution in ECG Signals, Although this 12-lead ECG device is not able to display all 12 leads simultaneously, it can display them sequentially [9, 21]. The last is research discussing the Design and Construction of a 12-lead Electrocardiogram

Device Developed on the ADS1293 Chip Platform. In future research it is recommended to be able to use a real-time frequency data recording system, meaning that all 12 lead inputs can be read simultaneously in the on state [22–24].

The main objective of this research is to propose the lowest sampling frequency and sampling resolution with the smallest bit depth. This will provide sufficient data representation to achieve accurate QRS detection The results of this study are to develop and construct a 12-lead electrocardiogram device that has a compact and user-friendly design and can communicate with a computer [25–28].

Based on the previous research that has been presented, the conclusion that can be drawn by the author is that, although there are many studies on electrocardiographs (ECG), there has been no development in making a 12-lead ECG by utilizing the ADS1293 CJMCU module that can be displayed on a personal computer (PC) [24, 29, 30]. In addition, the author analyzes the sampling rate of the STM32F7 microcontroller applied to the 12-lead ECG that has been designed [21, 31] The choice of using the STM32F7 microcontroller as the core of this research because the STM32F7 has a 32-bit resolution and can sample capacity reaching 2msps, the higher the resolution capability that can be achieved, so the measurement results become more accurate [32]. By using the method of taking precordial tapping data, namely V1–V6, and limb leads consisting of three bipolar leads (I, II, III) and three enhanced unipolar leads (aVR, aVL, aVF). The research has contributed to finding the optimal sampling rate for precordial leads and displaying 12 channels of ECG signals simultaneously.

2 Materials and Method

This research uses Phantom ECG as the data collection object. This study analyzes the use of the best sampling rate to record 12-lead ECG using the ADS1293 CJMCU module and processing with STM32F7. The sampling rate is set at nine values, namely 200 Hz, 300 Hz, 400 Hz, 500 Hz, 600 Hz, 700 Hz, 800 Hz, 900 Hz and 1000 Hz. Each sampling rate value is analyzed using the Signal to Noise Ratio (SNR) method to see the quality of the signal produced. The ADS1293 CJMCU module takes the lead value from the precordial and limb lead, then processes it with the STM32F7 and then the results are displayed on the PC.

Signals in the human heart which in this study used phantom ECG as a simulation were detected using cables connected to 12 sides of the heart including Lead I, Lead II, Lead III, aVR, aVL, aVf, V1, V2, V3, V4, V5 and V6 using RA, LA, RL, Leads LL, V1, V2, V3, V4, V5, and V6. Next, in the processing block, the signal that has been captured goes into the network buffer on the ADS1293 CJMCU module so that the signal detected is the actual signal and no signal is lost. After that, the output signal is connected to the Digital pin of the STM32F7 microcontroller. Then the data will be processed in the microcontroller, which is then implemented in the microcontroller program. This data will be displayed in the form of ECG signals in the software which is sent via USB connected from the device to the PC.

The ADS1293 Module is a circuit that has three bioelectric signal processing channels. Each module has six input pins that can be used. In the module, the first and second

pins are used as processing for lead 1, and the second and third pins are used for processing lead 2. Meanwhile, the other leads used are generated using the formula stated in the dataset as follows Eq. (1).

$$\text{Lead III} = \text{Lead II} - \text{Lead I} \tag{1}$$

The formula above is used to produce lead III. Lead III is the result of a comparison of lead readings between the left foot (LL) and the left hand (LA). The ADS1293 CJMCU module itself is accompanied by a conversion formula for its reading, making it easier for users to minimize the use of cables and circuits.

$$\text{Lead aVR} = -(\text{Lead I} + \text{Lead II})/2 \tag{2}$$

Lead aVR or enlarged Right Vector is a lead obtained by comparing the lead on the right hand (RA) with the Goldberger value 1 (G1). This value is obtained by dividing the Lead II and Lead I subtraction operations. The results of this operation are simplified by using the formula in Eq. (2).

$$\text{Lead aVL} = \text{Lead I} - \text{Lead II}/2 \tag{3}$$

Lead aVR or enlarged Right Vector is a lead obtained by comparing the lead on the left hand (LA) with the Goldberger 2 (G2) value. This value is obtained from the operation of reducing Lead I and Lead II which is divided by two. The results of this operation are simplified by using the formula in Eq. (3).

$$\text{Lead aVF} = \text{Lead II} - \text{Lead I}/2 \tag{4}$$

Lead aVR or enlarged Right Vector is the lead obtained by comparing the lead on the left leg (LL) with the Goldberger 3 (G3) value. This value is obtained from the operation of subtracting Lead II and Lead I which is divided by two. The results of this operation are simplified by using the formula in Eq. (4).

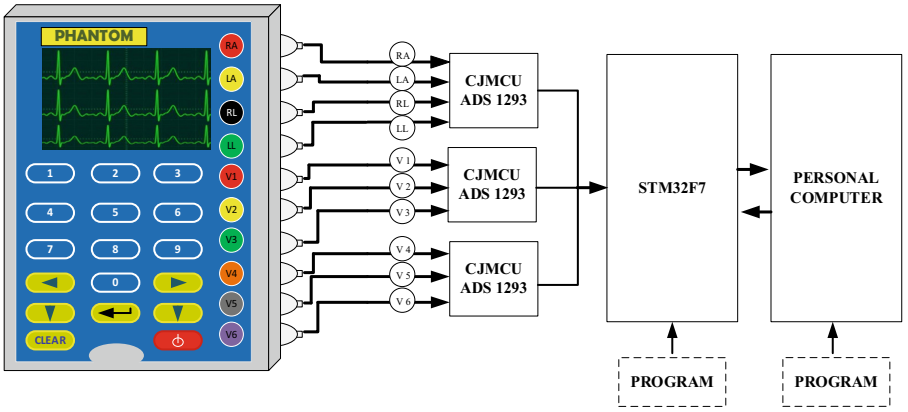


Fig. 1. Blog Diagram of Lead connection on Phantom ECG with ADS1293 CJMCU module

Figure 1 shows the data obtained from the phantom ECG simulation. We can set the BPM value and voltage of the signal that will be generated. The output from the phantom is given by electrode cables in the form of cables V1, V2, V3, V4, V5, and V6 as Prekordial Lead and Lead I, Lead II, Lead III, aVR, aVL, aVF as Limb Lead. The value from each point will be processed on the ADS1932 CJMCU ECG module, where three leads are processed on one module, so two modules are needed for precordial signal processing. From these two modules, it is converted into a signal on the STM32F7 which is then displayed on the personal computer via a USB cable.

As shown in Fig. 2, the system operates by initializing the program, which includes defining the pinout used as input. From program initialization, data is sent to the PC via USB cable, if the data is not sent, the process will remain at the initialization stage. The data received will be determined by the sampling rate used. Meanwhile, the program used to record data on the module is carried out by initializing the program on the STM32F7 via a USB cable. After initialization is carried out, the sampling rate is selected. Before the data is sent, the baud rate is selected. The data sent will be displayed on the personal computer in real-time.

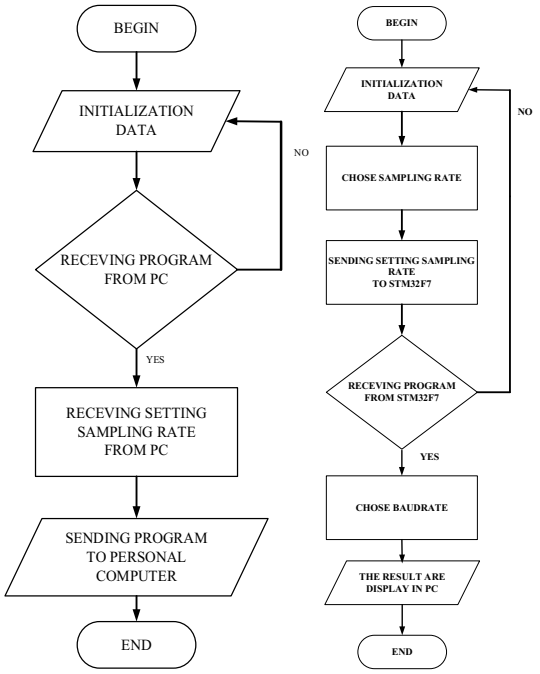


Fig. 2. The System Flowchart

The data will continue to be displayed in the Python program. The data is displayed in real-time as long as the sensor is still receiving voltage input. The sensor reading speed is influenced by the baud rate and specifications of the personal computer (PC) used. If the signal does not appear, the data will immediately be damaged and form ripples. In

the reading process, all data is carried out on the STM32F7 microcontroller, where the data is programmed in the Arduino IDE software, which after uploading will be read and stored. When the Python program is run, the data will immediately appear in the Python GUI.

2.1 Data Analysis

After taking many readings of a given parameter, the average value may be calculated using Eq. (5):

$$\overline{x} = \frac{x1 + x2 \ldots + xn}{n} \tag{5}$$

in this expression, x represents the average of n measurements, x1 represents the first measurement, x2 represents the second measurement, and so on up to xn. The standard deviation is a statistical measure of the dispersion of a group of numbers around the mean. The formula for the standard deviation (SD) is as follows in Eq. (6):

$$SD = \sqrt{\frac{\sum(xi - \overline{x})^2}{(n-1)}} \tag{6}$$

if n is the total number of measurements, then xi is the intended value distribution, and x is the weighted average of the measurements. As a kind of uncertainty, UA is present in all measurement results. The formula for the degree of uncertainty is given by Eq. (7):

$$UA = \frac{SD}{\sqrt{n}} \tag{7}$$

the sum measurement uncertainty, UA, is equal to the product of the standard deviation, SD, n. The term "deviation" refers to a result that is beyond the typical range of module measurements with the average measurement results on the measuring instrument. The deviation formula is shown in the Eq. (8):

$$\text{Deviation} = \overline{x}\text{module} - \overline{x}\text{measuring tool} \tag{8}$$

where x is an average value. An error occurred in the system. The standard deviation of the data is the lowest possible error value. Errors may point to problems with the model or design compared to the standard. The formula for the error is given in Eq. (9):

$$\%\ \text{error} = \frac{(xn - x)}{xn} \times 100\% \tag{9}$$

The measured design value is x, but the machine calibrator value is xn. These formulas are used to evaluate the signal results from the readings. The signal displayed is 24-bit data from the ADS1293 reading. The average results will be evaluated to determine what sampling rate is the most optimal. This formula will be implemented in the SNR calculation. The SNR results are processed in Mathlab using two formulas. From these results, the best sampling rate will be produced.

3 Result

In this research, there is one box that uses 3 modules that have been assembled. Figure 3 is a 12-lead 12-channel ECG, the circuit in the box consists of 3 ADS1293 CJMCU modules and an STM32F7 Nucleo 144 microcontroller using USB type b micro cable communication. The module is also given a jack which is connected to the electrode cable. The resource used to use the module is a USB cable connected to a personal computer (PC). The output from the electrode jack will become input from the ADS1293 CJMCU ECG module, which is installed on the PCB layout. This layout is also adapted to the microcontroller pins of the STM32F7.

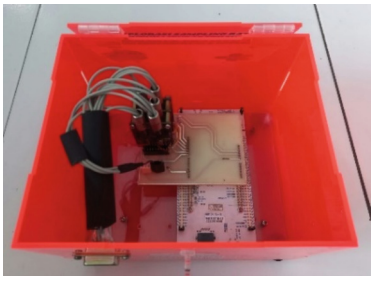


Fig. 3. The Design of a module

Figure 4 is a 12-channel display of the device that receives a 12-lead ECG signal. By displaying 12 channels simultaneously at the same period. The only way to use this display is to connect the USB type b connection from the microcontroller to the PC and open the display application then the 12-channel ECG display can be used. On the display, you can see the shape of the signal produced starting from the extremities and precordial. The voltage and time data values are also displayed in real-time, so we can calculate the Bits per minute (BPM) value of the patient.

The measurements taken were 12-lead ECG data or signals using a phantom ECG calibrator using different sampling rates. The sampling rate used is 200–1000 Hz. The difference in sampling rate can be replaced by changing the delay in the Arduino IDE program. The following is the sampling rate formula. From the data obtained, we can see that the greater the SNR value, the better the quality of the signal produced, and vice versa. Then the sampling rate getting bigger can also result in a delay that also increases in the incoming signal input. The results of applying SNR to the module can be seen in Table 1.

From the results of experiments with nine calculations, namely at a sampling rate of 200–1000, the best SNR was obtained at a frequency of 600 Hz. This frequency is quite high with the input signal being processed. The cut-off frequency for cardiac signals is 0.05–100 Hz. For further processing which requires a sampling rate where the requirement is twice the input frequency, then this result is satisfactory. However, with a frequency that is too high, it can produce a signal that has a delay in reading.

Table 1. SNR (Signal to Noise Ratio) in a 12-lead 12 Channel ECG model with several settings of Sampling Rate

Sampling frequency (Hz)	Signal-to-noise ratio (dB)	Delay (s)
200	− 2.8315	0.5
300	− 0.78819	1
400	12.8386	2
500	− 2.1051	3
600	15.997	4
700	− 3.1459	5
800	− 4	6
900	12.3749	7
1000	− 15.4391	8

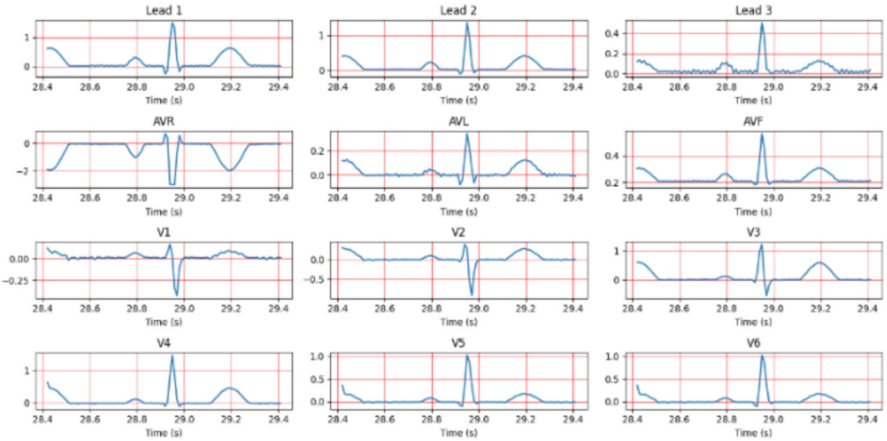


Fig. 4. The display of 12-lead 12 channel device on PC

4 Discussion

In exploring the use of sampling rate to obtain the most optimized signal on a 12-lead ECG, the overall system was evaluated using the STM32F7 Nucleo 144 microcontroller and the ADS1293 CJMCU ECG module. The STM32F7 Nucleo 144 was chosen as the main microcontroller for controlling and processing data from the ADS1293 CJMCU ECG module. To obtain signals from the precordial leads, the ADS1293 CJMCU module requires several aspects to generate the precordial leads. In Table 2 are some of the aspects of obtaining the precordial and limb leads. The act pin provides a voltage difference or can be said to be a reference for reading each precordial tapping point, because basically, tapping points on V1-V6 can be read if there is a reference voltage. For the limb lead, there is a reference from the provisions of each lead.

Table 2. Limb and precordial lead references

Type	Lead	Formula
Limb lead	Lead 1	LA–RA
	Lead 2	LL–RA
	Lead 3	Lead II–Lead I
	aVR	– (I + II) / 2
	aVL	I–II / 2
	aVF	II–I / 2
Prekordial lead	V1	V1–WCT
	V2	V2–WCT
	V3	V3–WCT
	V4	V4–WCT
	V5	V5–WCT
	V6	V6–WCT

WCT (Wilson Central Terminal) is the arithmetic mean of the voltages between the extremity electrodes (RA, LA, LL) at the same time. This helps correct for potential differences that may exist between the right and left sides of the body, allowing for a more accurate evaluation of the electrical activity of the heart. CJMCU ADS1293 module was used to record ECG data with the conversion of analog signals into digital format for further analysis. Various sampling rates were explored, including 200 Hz, 300 Hz, 400 Hz, 500 Hz, 600 Hz, 700 Hz, 800 Hz, 900 Hz and 1000 Hz, to determine the most optimal sampling rate setting.

In previous research, there were quite high delivery delays and lost data [5]. This research can anticipate this by determining the sampling rate of the ECG signal. With a small sampling rate, the resulting delay is quite low but this affects the imperfection of the resulting signal [33]. On the other hand, with a high sampling rate value, the signal results are better, this can be seen from the resulting Signal to Noise Ratio (SNR) value. However, the delay obtained is quite high. Because of this, to get good signal results with a small delay, observation and evaluation are needed using the Signal to Noise Ratio (SNR) method.

The weakness of this research is that the signal displayed is still affected by electromagnetic interference from outside and should be minimized using a digital filter and the data entering the PC display input still has a delay, and can still be improved using better devices or hardware.

The benefit of this research is to obtain the best sampling rate resolution when implementing a 12-lead ECG with a 12-channel display on a PC using the ADS 1293 CJMCU module with a microcontroller in the form of an STM32F7. By selecting a good sampling rate resolution, signal quality can be improved, making it easier to diagnose heart signals later.

5 Conclusion

The results show that of the various sampling rates explored (ranging from 200 Hz to 1000 Hz), a sampling rate of 400 Hz provides the most optimal signal resolution. The explanation for this result is that, while higher sampling rates can provide finer details in the signal, in this case, the 400Hz sampling rate provides a good balance of high enough resolution. In Table 1, we can see the SNR values from the sampling rate results of 200 Hz, 300 Hz, 400 Hz, 500 Hz, 600 Hz, 700 Hz, 800 Hz, 900 Hz, 1000 Hz, so the greater the sampling rate, the greater the delay on the display, then the greater the SNR value at each sampling rate, the better the signal quality compared to the noise in the signal. A sampling rate of 400 Hz has been shown to provide sufficient information for ECG analysis with sufficient accuracy. For further development, it is hoped that the module can reduce electromagnetic interference from outside and reduce program delays so that the ECG signal produced is also maximized.

References

1. Brucal SGE, Guemo JR, Montante SJB, Ng CAE, Wong DJC (2018) A development of a low-cost 12-lead electrocardiogram monitoring device using android-based smartphone. In: 2018 IEEE 7th Global Conference on Consumer Electronics GCCE 2018, pp 689–690 (2018). https://doi.org/10.1109/GCCE.2018.8574836
2. Darmawan RI, Surtono A, Apriyanto DK, Supriyanto A (2022) Design of computer based 12 lead ECG using STM32F401 microcontroller. J Energy, Mater Instrum Technol 3:147–156. https://doi.org/10.23960/jemit.v3i4.127
3. Brucal SGE, Aguirre JVCT, Macatangay SD, Rubia WU, Zamora AM (2018) Development of a 12-lead ECG signal processing algorithm using NI LabVIEW® and NI ELVIS®. In: 2018 IEEE 7th Global Conference on Consumer Electronics GCCE 2018, pp 186–189. https://doi.org/10.1109/GCCE.2018.8574627
4. Ye X, Lu Q (2020) Automatic classification of 12-lead ECG based on model fusion. In: Proceeding—2020 13th international congress image signal process. Biomedical engineering and informatics, CISP-BMEI 2020, pp 733–738. https://doi.org/10.1109/CISP-BMEI51763.2020.9263559
5. Irianto BG, Budhiaji B, Syaifudin S (2016) Design of electro cardiograph machine based on ATmega microcontroller. Indones J Electr Eng Comput Sci 2:328–333. https://doi.org/10.11591/ijeecs.v2.i2.pp328-333
6. Ionescu CM, Neckebroek M, Ghita M, Copot D (2021) An open source patient simulator for design and evaluation of computer based multiple drug dosing control for Anesthetic and hemodynamic variables. IEEE Access 9:8680–8694. https://doi.org/10.1109/ACCESS.2021.3049880
7. Ali I, Greifeneder F, Stamenkovic J, Neumann M, Notarnicola C (2015) Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. Remote Sens 7:16398–16421. https://doi.org/10.3390/RS71215841
8. Zhang J et al (2021) MLBF-net: a multi-lead-branch fusion network for multi-class arrhythmia classification using 12-Lead ECG. IEEE J Transl Eng Heal Med 9:1–11. https://doi.org/10.1109/JTEHM.2021.3064675
9. Tomasic I, Trobec R, Rashkovska A, Lindén M (2015) Impact of electrode misplacement on the accuracy of 12-lead ECG synthesized from differential leads. In: 2015 38th International Convention on Information and Communication Technology, MIPRO 2015—Proceeding, pp. 365–369 (2015). https://doi.org/10.1109/MIPRO.2015.7160296

10. Prabhakararao E, Dandapat S (2019) Automatic quality estimation of 12-lead ECG for remote healthcare monitoring systems. In: 2018 IEEE EMBS conference on biomedical engineering science IECBES 2018, Proceeding, pp 554–559 (2019). https://doi.org/10.1109/IECBES.2018.8626686
11. Tomasic I, Trobec R (2014) Electrocardiographic systems with reduced numbers of leads-synthesis of the 12-Lead ECG. IEEE Rev Biomed Eng 7:126–142. https://doi.org/10.1109/RBME.2013.2264282
12. Prabhakararao E, Dandapat S (2020) Attentive rnn-based network to fuse 12-lead ecg and clinical features for improved myocardial infarction diagnosis. IEEE Signal Process Lett 27:2029–2033. https://doi.org/10.1109/LSP.2020.3036314
13. Subramaniam SD, Dass B (2021) Automated nociceptive pain assessment using physiological signals and a hybrid deep learning network. IEEE Sens J 21:3335–3343. https://doi.org/10.1109/JSEN.2020.3023656
14. Gargiulo GD (2015) True unipolar ECG machine for wilson central terminal measurements. Biomed Res Int (2015). https://doi.org/10.1155/2015/586397
15. He K, Zhong G, Ding X, Yang C (2020) Recognition of premature ventricular contraction beat from 12Lead ECG based on a novel detection function of QRS onset. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, EMBS. 2020-July, pp 349–352. https://doi.org/10.1109/EMBC44109.2020.9175775
16. Zhang Q, Frick K (2019) All-ECG: a least-number of leads ECG monitor for standard 12-lead ECG Tracking during Motion. In: 2019 IEEE Healthc Innov Point Care Technol HI-POCT 2019, pp 103–106. https://doi.org/10.1109/HI-POCT45284.2019.8962742
17. Padhy S, Dandapat S (2016) Synthesis of 12-lead ECG from a reduced lead set using singular value decomposition. 2015 IEEE Power, Communication Information Technology Conference PCITC 2015—Proceeding, pp 316–320. https://doi.org/10.1109/PCITC.2015.7438183
18. Tabassum T, Ahmed M (2020) A simplified cardiac conduction model and twelve-lead ECG generation. In: 2020 International conference on computer, electrical & communication engineering, ICCECE 2020. https://doi.org/10.1109/ICCECE48148.2020.9223114
19. Adityaputra GB, Tasripan T, Sardjono TA (2019) Rancang Bangun Elektrokardiograf 12-Leads Untuk Sistem Pengawasan Kesehatan Jantung Jarak Jauh. J Tek ITS 8. https://doi.org/10.12962/j23373539.v8i1.38341
20. Gifari MW, Zakaria H, Mengko R (2015) Design of ECG homecare: 12-lead ECG acquisition using single channel ECG device developed on AD8232 analog front end. In: Proceedings of the 5th international conference on electrical engineering and informatics: bridging knowledge between academic and industrial community, ICEEI 2015, pp 371–376. https://doi.org/10.1109/ICEEI.2015.7352529
21. Zhang Y, Hou Z (2015) An algorithm for evaluating the ECG signal quality in 12 lead ECG monitoring system. In: Proceeding IEEE international conference on. software engineering and service science ICSESS. 2015-Novem, pp 453–456. https://doi.org/10.1109/ICSESS.2015.7339095
22. Ajdaraga E, Gusev M (2018) Analysis of sampling frequency and resolution in ECG signals. In: 2017 25th telecommunications Forum TELFOR, TELFOR 2017—Proceeding 2017-Janua, 1–4. https://doi.org/10.1109/TELFOR.2017.8249438
23. Ortiz KJP, Davalos JPO, Eusebio ES, Tucay DM (2018) IoT: electrocardiogram (ECG) monitoring system. Indones J Electr Eng Comput Sci 10:480–489. https://doi.org/10.11591/ijeecs.v10.i2.pp480-489
24. Kaewfoongrungsi P, Hormdee D (2017) The comparison between linear regression derivings of 12-lead ECG signals from 5-lead system and EASI-lead system. In: 2017 17th international symposium on communications and information technologies 2018-Janua, 1–6. https://doi.org/10.1109/ISCIT.2017.8261221

25. Nguyen TN, Duong TT, Omer H, Sulieman A, Bradley DA (2023) The design and construction of a 12-Channel electrocardiogram device developed on an ADS1293 Chip Platform. Electron 12. https://doi.org/10.3390/electronics12112389
26. Ionescu C, Lopes A, Copot D, Machado JAT, Bates JHT (2017) The role of fractional calculus in modeling biological phenomena: a review. Commun Nonlinear Sci Numer Simul 51:141–159. https://doi.org/10.1016/J.CNSNS.2017.04.001
27. Karlsson JS, Roeleveld K, Grönlund C, Holtermann A, Östlund N (2009) Signal processing of the surface electromyogram to gain insight into neuromuscular physiology. Philos Trans R Soc A Math Phys Eng Sci 367:337–356. https://doi.org/10.1098/RSTA.2008.0214
28. Ivanovic MD, Miletic M, Subotic I, Boljevic D (2019) Signal quality in reconstructed 12-lead ambulatory ECGs recorded using 3-lead device. In: Proceeding annual international conference of the IEEE engineering in medicine and biology society. EMBS, pp 5481–5487. https://doi.org/10.1109/EMBC.2019.8857251
29. Xu Y, Wang J, Ning X (2010) Characteristic parameters of multifractal singularity spectrum of synchronous 12-lead ECG signals. In: Proceeding—2010 3rd international conference on biomedical engineering and informatics, BMEI 2010, vol 2, pp 795–798. https://doi.org/10.1109/BMEI.2010.5639945
30. Joy E, Joseph RB, Lakshmi MB, Joseph W, Rajeswari M (2021) Recent survey on emotion recognition using physiological signals. In: 2021 7th international conference on advanced computing and communication ICACCS 2021. pp 1858–1863. https://doi.org/10.1109/ICACCS51430.2021.9441999
31. Chen YJ, Liu CL, Tseng VS, Hu YF, Chen SA (2019) Large-scale classification of 12-lead ECG with deep learning. In: 2019 IEEE EMBS international conference on biomedical health informatics, BHI 2019 Proceeding, pp 1–4. https://doi.org/10.1109/BHI.2019.8834468
32. Solieman H, Ripka D (2021) Convert standard ECG leads to orthogonal leads to obtain the heart vector magnitude. In: Proceeding 2021 IEEE conference of russian young researchers in electrical and electronic engineering ElConRus, pp 1839–1842. https://doi.org/10.1109/ElConRus51938.2021.9396159
33. Mulyadi IH, Nelmiawati SE (2018) Segment-specific EASI coefficients for improving accuracy of derived 12-lead electrocardiography. In: Proceeding of 2018 international conference on applied engineering, ICAE 2018, pp 1–4. https://doi.org/10.1109/INCAE.2018.8579385

Increasing the Accuracy of Body Weight Sensor Readings Using the Kalman Filter on the Smart Baby Scale for Monitoring Toddler Nutrition

Syaifudin(✉), Nur Latifatul Arifah, and Endro Yulianto

Department of Medical Electronics Technology, Poltekkes Kemenkes Surabaya, Surabaya, Indonesia
syaifudin_Lutfi@poltekkes-surabaya.ac.id

Abstract. Data from the Indonesian Nutritional Status Survey indicates that 21.6% of toddlers in the country suffer from stunting (SSGI). To help prevent stunting, anthropometric measurements are often conducted in health institutions to monitor child growth. This research aims to develop an anthropometric assessment technique equipped with a Kalman filter to reduce fluctuations and improve the accuracy of weight readings by load cell sensors. The tool incorporates an Internet of Things (IoT) based system that will automatically assess the nutritional condition of toddlers. This research contributes by integrating infant and toddler weight measurements into one compact tool, enhancing weight value accuracy, calculating z-score values automatically, and displaying a graph of weight growth compared to age and gender of infants and toddlers. The results anthropometric measurements will be displayed on Nextion device and forwarded to website, where visitors can download the data in Microsoft Excel format. Results indicated that error rate for toddler scales (standing) was 0.10%, which is lower than the average error rate for infant scales (lying) of 1.06%. The error value decreased from 0.122% before using the Kalman filter to 0.088%. The comparison between manual z-score calculation and system's calculation showed the system's z-score calculation was accurate, matching the manual calculation and correctly classifying the nutritional status of toddlers. This study has implication for assisttings health workers in monitoring toddler conditions and reducing errors in anthropometric data recording, potentially helping to reduce stunting cases. Additionally, the tool simplifies the weighing process of infants and toddlers in one compact device.

Keywords: Stunting · Anthropometric · Z-Score · Kalman Filter · Load Cell

1 Introduction

According to WHO estimates, by 2020 there will be 148.2 million cases of stunting in children under five years old worldwide [1]. Data from the Indonesian Nutrition Status Survey (SSGI) shows that 21.6% of children in Indonesia under the age of five will be stunted by 2022 [2, 3]. Malnutrition in children can be one of the factors causing stunting,

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 184–201, 2026.
https://doi.org/10.1007/978-981-95-9233-3_15

which is characterized by low height compared to age [4, 5]. In accordance with WHO guidelines, a child is considered stunted if the calculated z-score value is less than -2 SD on three measurements. Undernutrition increases the risk of death in children from common infections, worsens the frequency and severity of those infections, and impedes the recovery process. Inadequate nutrition during the first 1,000 days of a child's life can also result in stunted growth, which is difficult to correct and is associated with impaired cognitive function and reduced academic and professional performance [6–8].

Efforts to prevent stunting are carried out by monitoring the growth of toddlers such as by taking routine anthropometric measurements at health facilities [9–11]. Anthropometric measurements include measurements of body weight compared to age, and measurements of height/length which will be compared to age [12, 13]. Then the measurement data will be processed using z-score calculation analysis. Anthropometric indices and z-scores can be used to evaluate the nutritional condition of children under the age of five in developing countries. Z-score is a statistical measure that shows the position of the value of infant measurements (such as height, weight, and head circumference) from the average reference standard. Evaluation of the nutritional status and growth of infants based on the reference standards of percentile curves and Z-scores in anthropometric measures has been widely used to help assess nutritional status and growth, such as undernutrition (underweight, short and thin) and overnutrition (overweight and obese). In anthropometric calculations in toddlers, measurement results are calculated in percentile units (median). The Body Weight/Age parameter can be used for stunting detection to monitor the growth conditions of toddlers in very underweight, underweight, normal weight, or overweight risk [14, 15].

To obtain body weight measurement results, a load cell sensor is used which has a working principle by converting the received force or load into a measurable electrical signal. When using a load cell sensor, if the load measured on the loadcell moves, the data obtained will fluctuate and can affect the reading of body weight [13–16]. Therefore, a collaboration between readings from load cell sensors and digital filters is needed to get a more accurate weight reading value. Then a design is also needed from baby and toddler scales in one compact scale to make it easier for health workers to take anthropometric measurements on infants and toddlers. The classification of nutritional status values of infants and toddlers equipped with child growth charts can make it easier for health workers to identify and provide solutions regarding the health level of infants and toddlers. To achieve these targets, several researchers have made several studies on anthropometric tools. Research by designing an independent baby growth monitoring system with baby length and weight parameters which are then sent to an Android device using BT05 and then sent to the database [20]. However, the limitation of this research is that there is no classification of the baby's nutritional status value, so only the raw data from the measurements are displayed, and the measurement tool is specialized only for babies. Furthermore, research by producing a prototype of a baby weight and height measuring device using Arduino and displayed on the monitoring app system. This research can only display graphs of measurement results in the monitoring application [21]. Then, another study that aims to determine the value of nutritional status of toddlers based on the comparison of body mass index to age using Android smartphones and 20x4 LCD for display. This research is only devoted to toddlers under the age of

2 years and the classification of nutritional status values of toddlers is not complete, so it cannot determine the nutritional status value of obese toddlers [22]. In addition, there is a stunting early detection tool using anthropometric measurement methods with parameters used including weight, height, head circumference, and body temperature. However, the output data is printed using a thermal printer which is considered ineffective, because it can cause problems if the paper is lost, there is no classification of nutritional status values from toddlers, and the measurement tool can only be used for toddlers (standing) [10]. Previous research on the use of Kalman filters on loadcell sensors has also been conducted. Research in dynamic measurements to improve the accuracy of weight estimation on the scales. As a result, the use of Kalman filters has better performance in terms of standard deviation and completion time compared to the 5th order Butterworth filter [16]. In addition, in another study on the use of Kalman filters for other dynamic measurements, Kalman filters can be used to optimize results and reduce noise in measurements [23].

Based on the problem identification and references above, there is no infant and toddler weight measurement tool incorporated in one compact tool. So far, officers have entered toddler data, weight measurement results, and analyzed measurement results manually so that they are prone to errors. Several studies have determined the classification of nutritional status values of toddlers, but the classification of nutritional status values is only up to normal toddlers, unable to detect obese toddlers [21–23].ere is no output that displays the value of nutritional status and child growth charts to facilitate monitoring of toddler nutrition from month to month. And there is no digital filter added to the load cell sensor to increase the accuracy of the sensor reading value when the weighted load moves. Therefore, this research aims to create an automated method for measuring Body Weight/Age anthropometry in toddlers equipped with Internet of Things (IoT) based storage. With this automatic anthropometric measurement, parents (companions) no longer need to bring KMS (The Health Card) every time they do manual recording by officers or health workers [27]. Anthropometric measurement results will be processed and filtered by a microcontroller. Then the measurement results will be displayed on Nextion and will be sent to the website, where the data can be downloaded in Microsoft Excel format. So that this module is expected to help health workers in carrying out measurement activities and monitoring anthropometric results of infants and toddlers by minimizing data entry errors so that it can reduce stunting cases that occur in Indonesia. As well as simplifying the weighing of infants and toddlers in one complex tool.

2 Materials and Method

The Kalman Filter method is used to combine various sensor data and estimate the most likely value of a variable based on the available data. To predict a process, the Kalman filter employs feedback control. Prediction and update are the two main phases of the Kalman filter. The prediction equations are responsible for determining the estimated value for the next step. Meanwhile, the updated equations are responsible for incorporating new measurements into the estimate, thereby improving the accuracy of the final estimate. Thus, the output from using the Kalman filter is expected to reduce fluctuations

or noise in the signal or input data [18, 28]. The evolution of the state is described by Eq. (1):

$$X(t+1) = A(t)X(t) + b(t) + w(t), \tag{1}$$

where X is the state vector we intend to estimate, $A(t)$ is the known square transition matrix of the process. The control $b(t)$ is given, and there is a zero-mean process noise $w(t)$ with known covariance $r_{w(t)}$. The measured vector $y(t)$ is given by the measurement equation, as shown in Eq. (2):

$$y(t) = H(t)X(t) + v(t) \tag{2}$$

where $H(t)$ represents the rectangular measurement matrix, and $v(t)$ denotes the zero-mean measurement noise with a known covariance $r^v(t)$. The dimensions of $w(t)$ are the same as those of $X(t)$, while the dimensions of $v(t)$ correspond to those of $y(t)$. The covariance of the state vector $X(t)$ is

$$P(t) = E[(X(t) - E[X(t)\Big]) (X^T(t) - E\Big[X^T(t)])], \tag{3}$$

where X^T in Eq. (3) denotes the transpose (and possibly the conjugate) of X. The goal of the Kalman filter is to infer the vector $X(t)$ from $y(t)$ such that the covariance matrix of $X(t)$ has the smallest possible norm (its trace). The estimation process involves the following steps:

- Prediction of the state X(t):

$$X_{t+1/t} = A(t)X(t) + b(t); \tag{4}$$

This Eq. (4) predicts the state $X_{t+1/t}$ at the next time step t + 1. The current state X(t) evolves through the system's dynamics described by A(t). The term b(t) adds any external controls or inputs. This prediction gives us an estimate of the state before any new measurements are taken.

- Intermediate update of the state covariance matrix, incorporating the evolution described by the process transition (Eq. (5)):

$$P_{t+1/t} = A(t)P(t)A^T(t) + r^w(t); \tag{5}$$

The covariance $P_{t+1/t}$ in Eq. (5) reflects the uncertainty in the predicted state. The current uncertainty P(t) is adjusted by the system's dynamics through A(t) and its transpose $A^T(t)$. The term $R_{w(t)}$ accounts for the randomness or noise in the process. This step updates our confidence in the prediction.

- Computation of the optimal gain (Eq. (6)):

$$K_{t+1} = P_{t+1/t}H^T(t+1)\Big(H(t+1)P_{t+1/t}H^T(t+1) + r^v(t+1)\Big)^{-1} \tag{6}$$

The Kalman gain K_{t+1} in Eq. (6) determines how much the new measurement y(t + 1) should influence the state update. It balances the uncertainty in the predicted state $P_{t+1/t}$ with the measurement uncertainty $R_v(t + 1)$. This gain tells us how to combine the prediction and the new measurement for the best estimate.

- Update of the state covariance matrix, as shown in Eq. (7):

$$P(t+1) = P_{t+1/t} - P_{t+1/t}H^T(t+1) \left(H(t+1)P_{t+1/t}H^T(t+1) + r^v(t+1)\right)^{-1} H(t+1)P_{t+1/t}, \quad (7)$$

Alternatively, this can be expressed as Eq. (8):

$$P(t+1) = \left[I - K_{t+1}(t+1)\right]P_{t+1/t}; \quad (8)$$

After incorporating the new measurement, the covariance P(t + 1) is updated, as shown in Eq. (8). The identity matrix I ensures the dimensions match. The Kalman gain K_{t+1} and measurement matrix H(t + 1) reduce the uncertainty in the state estimate, reflecting the added confidence from the new measurement.

- Computation of the new estimate of the state (Eq. (9)):

$$X(t+1) = X_{t+1/t} + K_{t+1}\left[y(t+1) - H(t+1)X_{t+1/t}\right]. \quad (9)$$

The new state estimate X(t + 1) adjusts the predicted state $X_{t+1/t}$ using the difference between the actual measurement y(t + 1) and the predicted measurement $H(t + 1)X_{t+1/t}$. This difference is weighted by the Kalman gain K_{t+1}, providing the most accurate estimate of the state at time t + 1, as shown in Eq. (9).

The Kalman filter is commonly used in systems that continuously change in dynamic conditions. It is widely employed because it offers advantages such as the ability to eliminate noise and estimate future outcomes, thereby improving the accuracy of a variable's results [17, 29].

After the module was designed, it was tested in several ways, such as using calibrator scales with weights ranging from 1–20 kg, testing on toddlers aged 0–5 years, and analyzing measurement data with and without Kalman filters to reduce the standard deviation of measurement values. Testing with calibrator scales was conducted by placing the scales in the center of the cross-section of baby or toddler scales. Data collection on toddlers aged 2–4 years was performed five times, with the toddlers standing in the center of the cross-section of the scales. For the assessment of the effect of using Kalman filters, data was collected 30 times for each weight placed in the center of the cross-section of the toddler scales. This research utilizes a load cell sensor as a weight detector and an ESP32 microcontroller for data processing. The results can be displayed on a Nextion display and sent to a website, where they are stored in Microsoft Excel format.

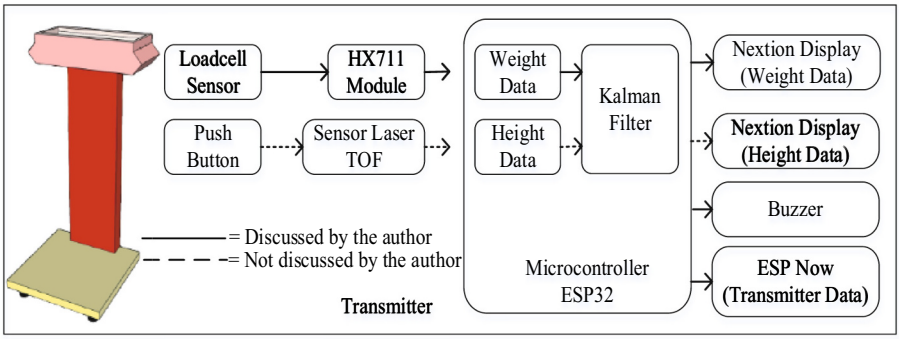


Fig. 1. Diagram Block of Transmitter. The sensor output is a weight value that will be processed by the Kalman filter in the ESP32. The weight value will be displayed on the Nextion display who sent via ESP Now**.**

This paragraph will explain Fig. 1 which is a transmitter block diagram. The input of this module is data from measuring body weight using a load cell sensor, so that an analog voltage is obtained which is then converted into digital data with the HX711 module. After obtaining the weight measurement results, the data will be processed on the ESP32 microcontroller and then the sensor data processing will be carried out using the Kalman filter method which functions to increase accuracy and reduce sensor reading noise. The results will be displayed on the Nextion display module (receiver) using ESP Now serial communication as a data transmitter marked by a buzzer indicator sounding.

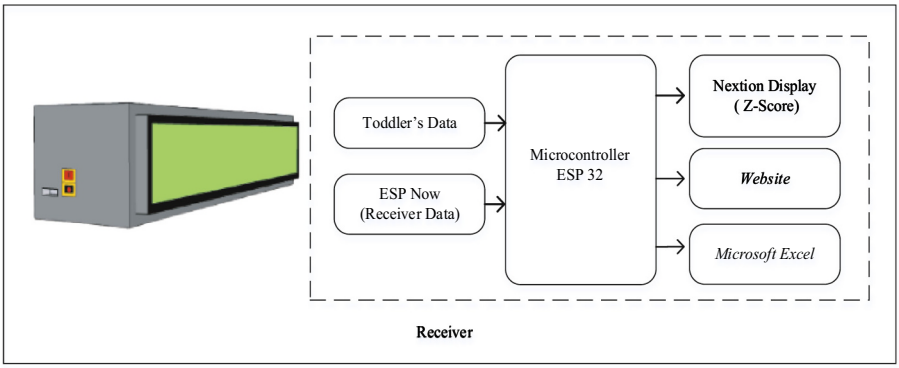


Fig. 2. Diagram Block of Receiver. Data input from the ESP Now Transmitter will appear on the nextion display.

Figure 2 is a block diagram of the display module that acts as a receiver. The display module gets input from toddler data and measurement modules via ESP Now serial communication. Toddler data entered such as toddler name, toddler age, parent name, and address. The data received from the measurement module is the reading of body weight and body length which is then corrected to get the correct reading before being sent to the website. On the website, the z-score calculation will be done. The display

on the website includes toddler identity, weight measurement results, height, toddler nutritional status, and toddler growth charts. The data can be downloaded in Microsoft Excel format.

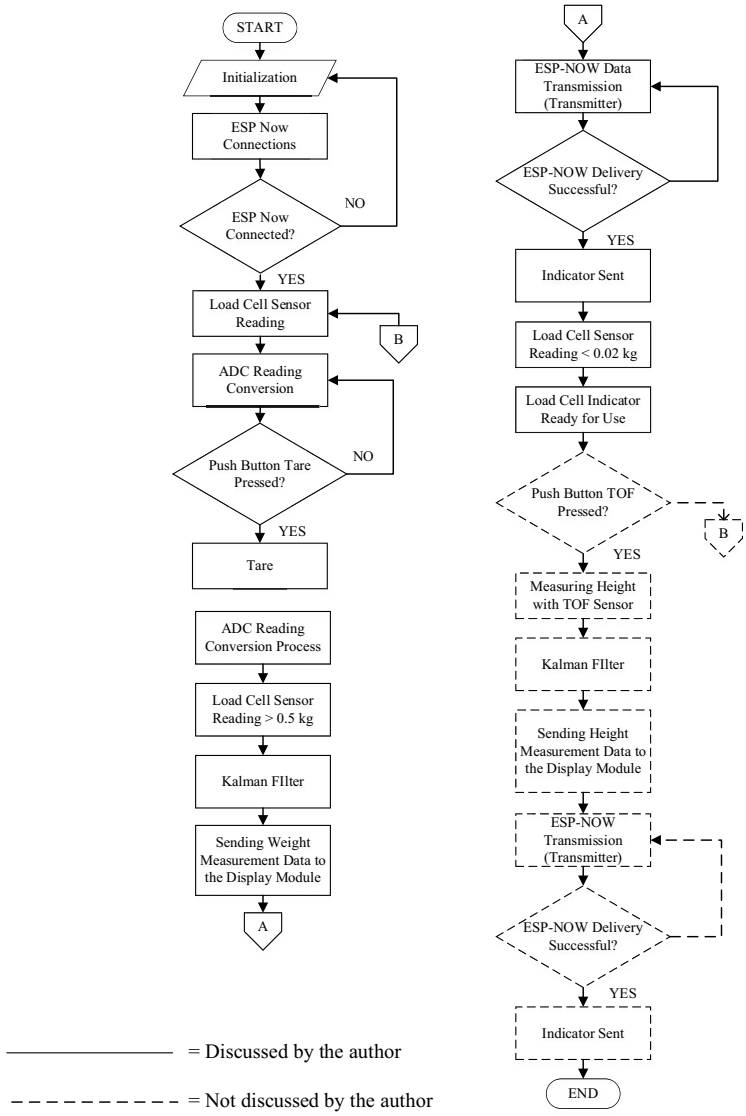


Fig. 3. Flowchart of the Transmitter: The final body weight data will be sent through the ESP-NOW receiver.

Figure 3 is a system Flowchart of the Transmitter. The process begins with initialization, followed by ESP-NOW serial communication between the transmitter and receiver. If the communication is not yet established, the connection process will continue to repeat until a successful connection is made. Once the ESP-NOW communication is connected between the transmitter and receiver, the load cell sensor begins reading the weight input. The weight data is then processed using the ADC reading conversion via the HX711 module. When the tare push button is pressed, the system will tare the current load, effectively resetting the scale to zero. After taring, the load cell sensor will continue reading the weight and converting the ADC reading using the HX711 module. If the weight exceeds 0.5 kg, the system starts counting, and the data from the load cell sensor is processed using a Kalman filter to ensure the resulting value is accurate and stable. The processed data is then transmitted to the display module (receiver) via ESP-NOW serial communication, indicated by a buzzer sound. Once the reading is complete and the load is removed, when the load cell sensor detects a weight of 0.02 kg or less, the buzzer will sound again to indicate that the load cell is ready for the next measurement. Finally, the data is sent to the display module (receiver) using ESP-NOW serial communication, with the buzzer signaling the successful transmission.

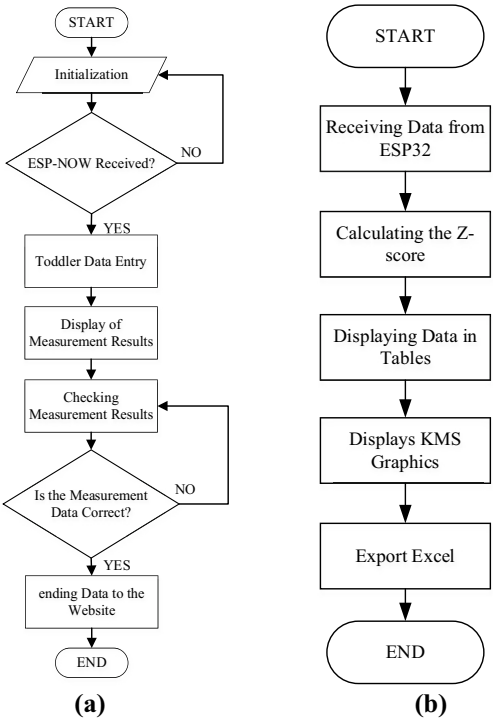


Fig. 4. Flowchart of Communication Serial. **(a)** Flowchart of Receiver **(b)** Flowchart of Website.

Refer to Fig. 4(a) which shows a system flow diagram of the display module that functions as a receiver. The process begins with data initialization. Next, toddler data can

be input on the Nextion display, and the ESP32 will read this data. If the ESP32 does not receive serial data from the measurement module (transmitter), the initialization process will repeat. Once the ESP32 receives the serial data, the display module will show the weight measurement data, which the user must validate. If an error is detected in the data, the retrieval process will be repeated. When the data is confirmed as correct, it will be sent to the website.

Based on Fig. 4(b) shows a website flowchart. The process starts with the website receiving data from the ESP32. The received data is then processed by calculating the z-score to determine the toddler's nutritional status. The z-score values are graphed to facilitate monitoring the toddler's nutritional status from month to month. Data stored on the website can also be downloaded in Microsoft Excel format.

Fig. 5. Measuring Body Weight in Toddler Respondents by the age under 5 years

Figure 5 is a picture of data collection on weight values from toddlers aged 48 months. When weighing, the toddler is in the middle position of the scale. Then the result of this toddler is risk of overweight. This body weight data collection was carried out 5 times for each toddler.

2.1 Data Analysis

Z-Score, is a statistical measure that indicates the position of an infant's measurements (e.g. height, weight, head circumference) from the mean of a reference standard. Evaluation of infant nutritional status and growth based on the reference standard of percentile curves and Z-scores in anthropometric measures has been widely used to help assess nutritional status and growth, such as undernutrition (underweight, short and thin) and overnutrition (overweight and obese). The formula to calculate the z-score value is as follows Eq. (10) [15]

$$\text{Z}-\text{Score} = \frac{(\text{actual anthropometric value} - \text{median reference value})}{\text{Standard deviation of reference}} \tag{10}$$

The reference data used in the z-score calculation process is based on data from the Regulation of the Ministry of Health of the Republic of Indonesia Number 2 of

2020 concerning Child Anthropometric Standards by comparing the body weight/age of toddlers with the average body weight/age of the population of toddlers of the same age and sex [30].

After obtaining the results of the z-score calculation, information on a toddler's nutritional status and growth can be obtained. A Z-score value lower than -2 indicates that the toddler is at risk of stunting or malnutrition, so special treatment is needed to improve the toddler's nutritional status. Meanwhile, a z score higher than + 2 indicates that the toddler has excessive growth or is overweight.

After taking multiple measurements of a specific parameter, the average value can be calculated using Eq. (11).

$$\bar{x} = \frac{x1 + x2 \ldots + xn}{n} \tag{11}$$

In this formula, the first measurement is denoted by x1, the second by x2, and so on up to xn. X represents the average of n measurements. A statistical indicator of a set of numbers' dispersion around the mean is the standard deviation. Equation is the formula for the standard deviation (SD), as shown in Eq. (12):

$$SD = \sqrt{\frac{\sum(xi - \bar{x})^2}{(n - 1)}} \tag{12}$$

where x is the average. A system error transpired. The data's standard deviation represents the lowest error value that can exist. Errors in comparison to the standard may indicate issues with the model or design. Equation of error's Eq. (13):

$$\%\ \text{error} = \frac{(xn - x)}{xn} \times 100\% \tag{13}$$

3 Result

The weight of babies and toddlers will be measured using the same module. For toddlers, the weight is measured at the bottom section where they stand. In contrast, for babies, the measurement is taken in the upper section where they lie down. The weight data obtained from the module for both babies and toddlers will then be pro-cessed by the microcontroller. The design of the baby and toddler weight measure-ment module is as follows:

Based on Fig. 6(a) the image shows the front view of the module that has been created. The upper section is designed for measuring the weight and length of babies under 2 years old or those who are lying down. The lower section is used to measure the weight of toddlers who can stand, typically aged 2 to 5 years. Thus, both babies and toddlers can be weighed using the same module.

Refer to Fig. 6(b) which shows an image of the front of the display module that has been created. The front panel includes a Nextion touchscreen for data input and a display for viewing the measurement results. On the side, there is a battery indicator. Inside the module, a circuit connects the ESP32 with the Nextion display.

(a) (b)

Fig. 6. Design results of the tool: (a) Baby and Toddler Weight Module: The top section is a scale for babies, while the bottom section is a scale for toddlers. (b) Display Module acting as a receiver on the Nextion display.

Results of data collection on toddler scales, the process of taking data on toddler scales or standing scales is carried out using calibrators in the form of weights that have different weight variation values. The purpose of the test is to find out whether the value read on the scale module is in accordance with the weight of the scales used. Data collection was carried out 6 times on 22 parameters. The following is the data on the toddler scales:

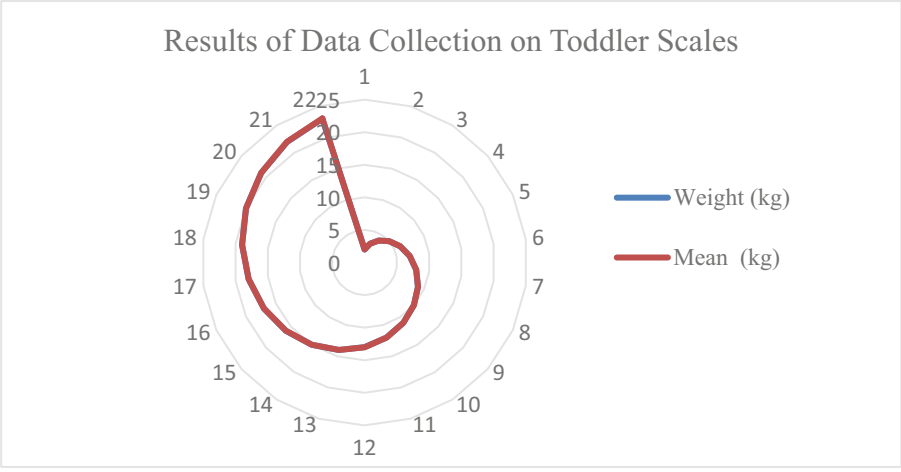


Fig. 7. Results of Data Collection on Baby Scales

Based on the results shown in Fig. 7, testing the scale module on a standing scale displays the average value of the measurement of 22 different weights. From the average value, a high error value is obtained at a weight of 2 kg with an error value of 0.16%. Then the lowest error value is at a weight of 4 kg by 0%. The average error value for toddler scales is 0.10%.

Results of data collection on baby scales, the data collection process on baby scales or sleep scales is carried out using a calibrator in the form of scales that have different

weight variation values. The purpose of the test is to find out whether the value read on the scale module matches the weight of the scales used. Data collection was carried out 6 times on 10 parameters. The following is the data on baby scales:

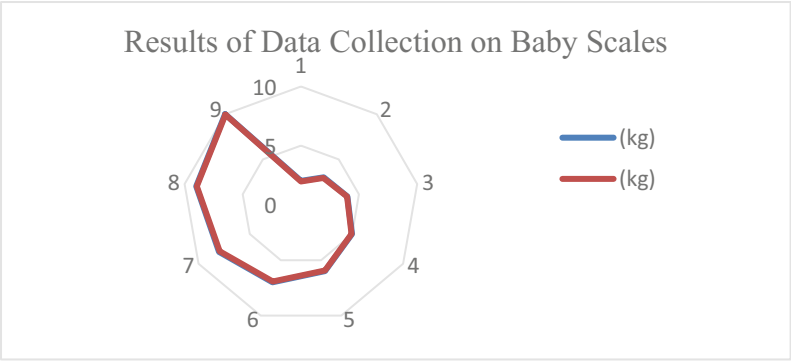


Fig. 8. Results of Data Collection on Baby Scales

Based on the results shown in Fig. 8, testing the scale module on a standing scale displays the average value of measurements from 10 different weights. From the average value, a high error value is obtained at a weight of 2 kg, with an error value of 1.83%. Then, the lowest error value is at a weight of 9 kg, by 0.57%. The average error value for toddler scales is 1.06%.

Measurements were carried out using 5 load variations and data was collected 30 times which were then calculated the average, the value of the standard deviation, and also the average standard deviation before and after the Kalman filter was applied.

Table 1. Results of Load Cell Sensor Measurement Analysis of Kalman Filter

Output	SD	Error (%)
Load Cell	0.004	0.122
Kalman Filter	0.002	0.088

Based on Table 1 we can see the comparison results when the Kalman filter is used on the load cell sensor readings and when the Kalman filter is not used. The standard deviation value before the Kalman filter is 0.004, then after being subjected to the Kalman filter, the standard deviation value or the distribution of data variations decreases or decreases to 0.002. Then the error value also decreases when before the Kalman filter is applied by 0.122% but after applying the Kalman filter the error value decreases to 0.088%. The following is a comparison graph of sensor output before and after Kalman filter.

In collecting data on toddlers directly, data collection on weight measurements was carried out on 4 toddlers aged 0–5 years with 5 repetitions of measurements on each toddler.

Table 2. Results of Weight Measurement and Nutritional Status of Toddlers

Respondent	Gender	Measurement Results			
		Age (Month)	Weight After Kalman Filter (kg)	Z-score	Nutritional Status
1	Male	48	20.77	2.35	Risk of Overweight
2	Female	33	10.79	−1.394	Normal
3	Male	26	12.22	−0.21	Normal
4	Female	23	7.08	−2.14	Under Weight

Based on Table 2, the module can automatically measure body weight and calculate the z-score value to determine the nutritional status of toddlers. Respondent 1 is at risk of being overweight, respondents 2 and 3 have normal nutritional status, and respondent 4 has a nutritional status value of less.

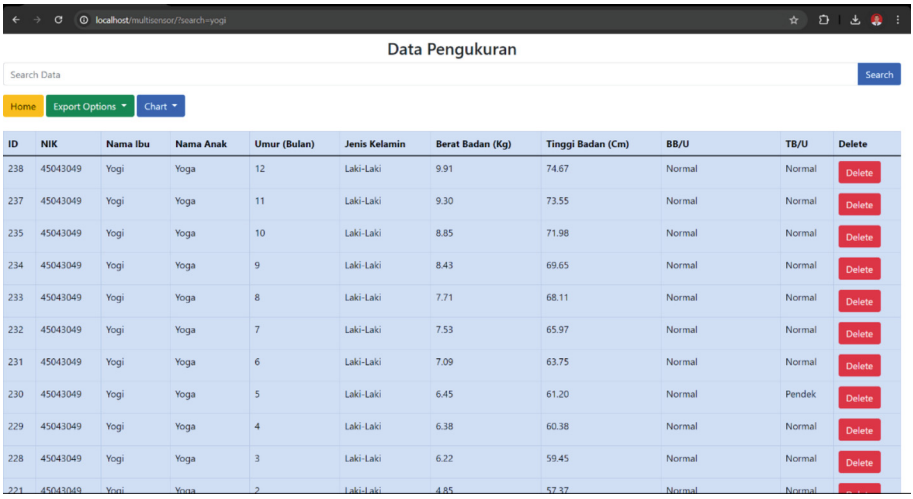


Fig. 9. Display on the Loacal Host Website. Identity of infants and toddlers, Anthropometric Measurements, and Nutritional Status Values

Refer to Fig. 9 is an image of the localhost website display in this study. On this website, there are some personal data of infants and toddlers such as ID which is the identity number of the baby or toddler who is weighing, then NIK is the Population Identification Number of the baby or toddler's mother, Nama Ibu is the mother's name,

Nama Anak is the name of the baby or toddler who is weighing, Umur is the age in months, Jenis Kelamin is gender, Berat (kg) is the weight value of the baby or toddler who has been weighed, Tinggi (cm) is the height of the baby or toddler, BB/U is the nutritional status value of toddlers based on z-score calculations that have been carried out by the system based on weight values compared to age and gender, TB/U is the nutritional status value of toddlers based on z-score calculations that have been carried out by the system based on height values compared to age and gender.

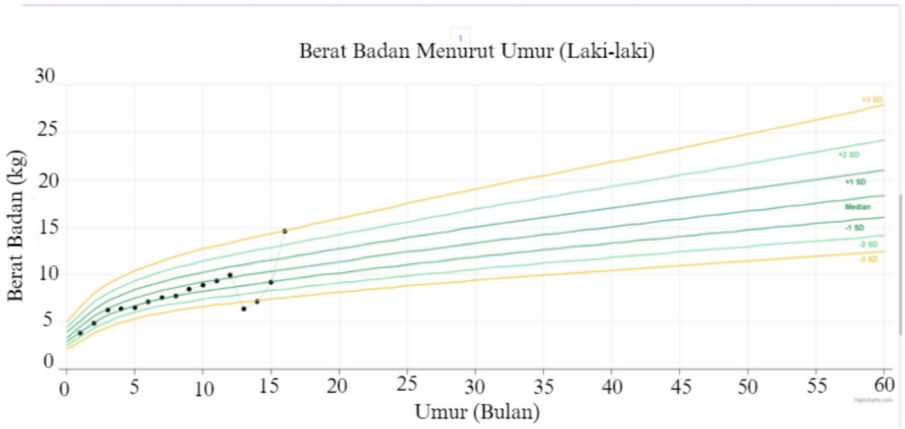


Fig. 10. Child Growth Graph on the Website Display. Month-to-Month Z-Score Results are Graphed.

Refer to Fig. 10 child growth chart on the website display is a way to make it easier to monitor the nutritional status of infants and toddlers. In the title Weight According to Age (Boys) is an explanation that the graph below is the nutritional status value of the z-score value based on weight and age in boys. Then on the x-axis is age (months) which means that on the x-axis is the value for the month (time) when the weighing was carried out. Then on the y-axis is Weight (kg) which means the weight value of the baby or toddler when weighing was carried out. so that the black dot at the x and y coordinates is the point where the weight of the baby or toddler is based on time to time each month. If the black dot value is below -2SD, this indicates that the baby or toddler is experiencing stunting, so a solution is needed to improve it.

Refer to Fig. 11 is a display in excel format downloaded from the website. Nama Anak is the name of the baby or toddler who was weighed, then in Umur (bulan) is the age of the baby or toddler, Jenis Kelamin is the gender of the baby or toddler, Berat Badan (kg) is the value of anthropometric measurements using a scale, Tinggi Badan (cm) is the height value of the baby or toddler, BB/U is the nutritional status value of toddlers based on z-score calculations that have been carried out by the system based on weight values compared to age and gender, TB/U is the nutritional status value of toddlers based on z-score calculations that have been carried out by the system based on height values compared to age and gender.

	D	E	F	G	H	I	J
1	Nama Anak	Umur (Bulan)	Jenis Kelamin	Berat Badan (Kg)	Tinggi Badan (Cm)	BB/U	TB/U
2	Yoga	12	Laki-Laki	9.91	74.67	Normal	Normal
3	Yoga	11	Laki-Laki	9.3	73.55	Normal	Normal
4	Yoga	10	Laki-Laki	8.85	71.98	Normal	Normal
5	Yoga	9	Laki-Laki	8.43	69.65	Normal	Normal
6	Yoga	8	Laki-Laki	7.71	68.11	Normal	Normal
7	Yoga	7	Laki-Laki	7.53	65.97	Normal	Normal
8	Yoga	6	Laki-Laki	7.09	63.75	Normal	Normal
9	Yoga	5	Laki-Laki	6.45	61.2	Normal	Pendek
10	Yoga	4	Laki-Laki	6.38	60.38	Normal	Normal
11							
12							

Fig. 11. Microsoft Excel Document View of Downloaded Data on the Website

4 Discussion

The analysis carried out to determine the accuracy of the load cell sensor is to analyze the measurement of the load cell sensor using a scale calibrator that has varying weight values and data collection is carried out 6 times on each load variation. The results will be compared when using toddler and baby scales. The results obtained are the highest error value on the toddler scales of 0.33% and the lowest error of 0% with an average error value of 0.10%. Then on the baby scales the highest error value is 1.83% and the lowest error is 0.57% with an average error value of 1.06%. Analysis is also carried out by comparing the value of measurement results when using a Kalman filter and sensor results without filtering with a Kalman filter. It is known that the use of the Kalman filter can reduce fluctuations in sensor reading values as indicated by a decrease in the standard deviation value before and after using the Kalman filter. The average standard deviation is reduced from 0.004 to 0.002. So that the distribution of data in the system becomes narrower. And with automatic calculation by the system, the results of the z-score value along with the nutritional status value of toddlers become more accurate. This is due to the minimal error caused by human calculation errors.

In research on anthropometric tools for early detection of stunting using parameters of weight, height, head circumference and body temperature, it has not been able to classify the nutritional status values of toddlers and the raw data output is still in the form of thermal printer paper which is prone to loss and damage [10]. Then in previous research that designed a monitoring system, the growth and development of children independently with parameters of body length and baby weight, also has not been able to classify the value of the nutritional status of babies based on data from measuring the weight and length of babies. Then in this study also only focuses on anthropometric measurements for babies only (sleeping position) [20]. So that compared to previous research, has a better error value in highest.

This research can combine anthropometric measurements of infants and toddlers in one compact tool, increase the accuracy of the results of reading the weight value with the Kalman filter method added in reading the value of the load cell, produce a classification of the value of the nutritional status of infants and toddlers obtained from weight readings, as well as inputting personal data and the results of measurement data

that will be recorded by the system and can be monitored via a website with data display and child growth charts which can then be downloaded in Microsoft excel format.

The weakness of this module is that when weighing the load weighed on the top scale or baby scale, the scale will sway slightly due to the less sturdy mechanical mechanism. In addition, when the weighed load exceeds the toddler weight standard, the module will sway slightly due to the placement of the load cell sensor which is not at an angle at each corner of the cross section at the bottom. Therefore, if the distance between the load cells is changed, the value generated by the load cell sensor becomes accurate.

It is hoped that this anthropometric tool for infants and toddlers can help health workers in measuring and monitoring the growth of infants and toddlers by minimizing data entry errors. Then with this tool health workers can also easily analyze the nutritional status values of infants and toddlers because there is no need to manually calculate the z-score value and draw the value of the child's growth chart manually, all of which have been done by the system. And with this tool, guardians of infants and toddlers no longer need to bring mother and child cards, because measurement data along with the nutritional status of toddlers can be seen on the website and can be downloaded in Microsoft Excel format. So that with this tool can facilitate health workers and reduce the incidence of stunting in Indonesia.

5 Conclusion

This research aims to increase the accuracy of reading load cell sensor values using the Kalman filter on the infant and toddler scale. It can be concluded that the baby and toddler scales have been successfully developed and can be used well. The resulting weight value before the Kalman filter was applied had a standard deviation value of 0.004 and decreased when the Kalman filter was applied by 0.002. Based on this data, the use of the Kalman filter on the load cell sensor does not have much effect in reducing the standard deviation value because the load cell sensor used for this scale has values that are quite stable and accurate. The Excel data downloaded from the website was very helpful for Kader Posyandu, who had previously recorded manually. For future research, this research could be improved by considering the creation of safe mechanical designs that support and stabilize the weighing process. This design should not affect baby and toddler weight readings. Additionally, it is important to position the load cell points carefully to ensure optimal sensor readings throughout the mechanical structure. Future developments could also include implementing a more user-friendly data input system using RFID technology and creating a mapping system to track the prevalence of stunting in various regions.

References

1. UNICEF/WHO/WORLD BANK (2021) Levels and trends in child malnutrition UNICEF/WHO/world bank group joint child malnutrition estimates key findings of the 2021 edition. World Health Organization, pp 1–32
2. Prendergast AJ, Humphrey JH (2014) The stunting syndrome in developing countries. Paediatr Int Child Health. 34:250–265

3. Beal T, Tumilowicz A, Sutrisna A, Izwardy D, Neufeld LM (2018) A review of child stunting determinants in Indonesia. Matern Child Nutr 14:e12617
4. Vaidya AS, Makkena G, Srihari V, Srinivas MB, Rao SK (2013) A sustainable solution for monitoring malnutrition in children in developing countries. In: 2013 IEEE global humanitarian technology conference: south asia satellite (GHTC-SAS). IEEE, pp 170–174
5. Mahmudiono T (2016) Child stunting in households with double burden of malnutrition: applications of behavioral epidemiology. Kansas State University
6. Hoque MA, Sayeed MA, Ahsan MR, Al Mamun MA, Salim F (2016) Nutritional status among under-5 Children of a selected slum in Dhaka city. Northern Int Med College J 7:143–145
7. Dewey KG, Begum K (2011) Long-term consequences of stunting in early life. Matern Child Nutr 7:5–18
8. Soliman A, De Sanctis V, Alaaraj N, Ahmed S, Alyafei F, Hamed N, Soliman N (2021) Early and long-term consequences of nutritional stunting: from childhood to adulthood. Acta Bio Medica: Atenei Parmensis 92
9. Wang Y, Chen H-J (2012) Use of percentiles and z-scores in anthropometry. Handbook of anthropometry: physical measures of human form in health and disease, pp 29–48
10. Umiatin U, Indrasari W, Taryudi T, Dendi AF (2022) Development of a multisensor-based non-contact anthropometric system for early stunting detection. J Sens Actuator Netw 11:69
11. Haile D, Nigatu D, Gashaw K, Demelash H (2016) Height for age z score and cognitive function are associated with Academic performance among school children aged 8–11 years old. Archives of Public Health. 74:1–7
12. Jorge CB, Doris CP (2014) Z-score anthropometric indicators derived from NCHS-1977, CDC-2000 and WHO-2006 in children under 5 years in central area of Peru. Univers J Public Health 2:73–81
13. Putri RA, Sendari S, Widiyaningtyas T (2018) Classification of toddler nutrition status with anthropometry calculation using Naïve Bayes Algorithm. In: 2018 International conference on sustainable information engineering and technology (SIET). IEEE, pp 66–70
14. Putri TE, Subagio RT, Sobiki P (2020) Classification system of toddler nutrition status using naïve bayes classifier based on Z-score value and anthropometry index. In: Journal of Physics: Conference Series. IOP Publishing, p 012005
15. Dibley MJ, Staehling N, Nieburg P, Trowbridge FL (1987) Interpretation of Z-score anthropometric indicators derived from the international growth reference. Am J Clin Nutr 46:749–762
16. Pitawala S (2020) State estimation for dynamic weighing using Kalman filter. In: Journal of Physics: Conference Series. IOP Publishing, p 012017
17. Welch GF (2020) Kalman filter. Computer Vision: A Reference Guide, pp 1–3
18. Ma'arif A, Iswanto I, Nuryono AA, Alfian RI (2019) Kalman filter for noise reducer on sensor readings. Sig Image Proc Lett 1:50–61
19. Muller I, de Brito RM, Pereira CE, Brusamarello V (2010) Load cells in force sensing analysis–theory and a novel application. IEEE Instrum Meas Mag 13:15–19
20. Qian H, Liu J, Wu Y (2019) A self-service scheme of infant scale for height and weight. In: 2019 IEEE MTT-S international microwave biomedical conference (IMBioC). IEEE, pp 1–3
21. Umiatin, Erlandita SM, Indrasari W (2019) Design baby mass and height monitoring system based on Arduino and Android application. In: AIP conference proceedings, p 30013. AIP Publishing LLC (2019)
22. Ardianto ET, Elisanti AD, Husin H (2022) Arduino and android-based anthropometric detection tools for Indonesian children. In: 2nd international conference on social science, humanity and public health (ICOSHIP 2021). Atlantis Press, pp 254–259
23. Halimic M, Balachandran W (1995) Kalman filter for dynamic weighing system. In: 1995 Proceedings of the IEEE international symposium on industrial electronics. IEEEm pp 786–791

24. De Onis M et al (2019) Prevalence thresholds for wasting, overweight and stunting in children under 5 years. Public Health Nutr 22:175–179
25. Das S, Alam MA, Mahfuz M, El Arifeen S, Ahmed T (2019) Relative contributions of the correlates of stunting in explaining the mean length-for-age z-score difference between 24-month-old stunted and non-stunted children living in a slum of Dhaka, Bangladesh: results from a decomposition analysis. BMJ Open 9:e025439
26. Rachmi CN, Agho KE, Li M, Baur LA (2016) Stunting, underweight and overweight in children aged 2.0–4.9 years in Indonesia: prevalence trends and associated risk factors. PLoS One. 11:e0154756
27. Organization WH (2019) Recommendations for data collection, analysis and reporting on anthropometric indicators in children under 5 years old
28. Le Roux J (1960) An introduction to Kalman filter. ASME J Basic Eng 82:34–45
29. Buchstaller D, Liu J, French M (2021) The deterministic interpretation of the Kalman filter. Int J Control 94:3226–3236
30. Ministry of Health of the Republic of Indonesia: Standar Antropometri Anak. Child Anthropometry Standards 21:1–9

Skin Cancer Identification Based on Dermoscopy Images Using Fractal and Gray Level Co-occurrence Matrix Features

Emmanuella Aurelia Rachel Passa[1], Endah Purwanti[1,3(✉)], Syifa Candiki Samatha[1,2], Alfian Pramudita Putra[1,3], and Rinci Kembang Hapsari[4]

[1] Biomedical Engineering Study Program, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia
endah-p-1@fst.unair.ac.id
[2] Biomedical Engineering Master Study Program, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia
[3] Biomedical Engineering Innovation Study Group, Faculty of Science and Technology, Airlangga University, Surabaya, Indonesia
[4] Department of Informatics Engineering Institute, Teknologi Adhi Tama, Surabaya, Indonesia

Abstract. The skin is the largest organ of the body and functions as a biologically active barrier. However the threat of disease is unavoidable, one of which is skin cancer. The main causative factor for skin cancer is ultraviolet radiation which is a strong carcinogen. Indonesian health profile reported that expert medical numbers are unevenly distributed, particularly in rural areas with substandard healthcare facilities. This is inversely proportional to the number of skin cancer cases especially in countries with high levels of sun exposure. The application of artificial intelligence technology may prove to be a significant breakthrough, potentially assisting medical personnel in providing a second opinion. Through this study, an artificial intelligence system is designed using the Support Vector Machine (SVM) algorithm that will assist medical personnel in identifying the type of skin cancer. The performance of the SVM system will be evaluated using the evaluation parameters to see which kernel provides optimal results in identifying image data of skin cancer patients. The study begins with data acquisition and processing used to improve desmoscopy image quality. The combination of fractal and Gray Level Co-occurrence Matrix (GLCM) is used as feature extraction and Support Vector Machine (SVM) algorithm as identification algorithm. Based on the research results, the system performance using only the fractal method does not provide optimal results. Thus, with the addition methods to the preprocessing and feature extraction stages to optimize its performance obtained accuracy results in the polynomial kernel of 97.71% for training and 100% for testing.

Keywords: Skin Cancer · Dermoscopy · Fractal · GLCM · SVM

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 202–218, 2026.
https://doi.org/10.1007/978-981-95-9233-3_16

1 Introduction

Skin is the largest organ of the body and serves as a biologically active barrier that metabolically separates internal homeostasis from the external environment [1]. It can protect human internal organs from heat, sunlight, dust, smoke, harmful microorganisms, infections around the body, and so on. Having various functions does not make the skin avoid the threat of disease, one of the dangerous diseases that attack the skin is skin cancer. Skin cancer is a dangerous type of cancer because it can cause death, disability, and a high financial burden [2]. Skin cancer is divided into melanoma and non-melanoma skin cancers (NMSCs) that are divided into Basal Cell Carcinomas (BCC) and Squamous Cell Carcinomas (SCC). It is also the most commonly diagnosed cancer group worldwide, with an estimated more than 1.5 million new cases in 2020 [3].

Skin cancer is a multifactorial skin disease, which increases the rate of skin cancer. Sunlight is a continuous spectrum of electromagnetic radiation that is divided into three main spectrums of wavelength, witch ultraviolet, visible light, and infrared [4]. Ultraviolet Radiation (UVR) can cause damage and cell changes in immunological function. This carcinogen can suppress the body's performance that plays an important role in preventing premature aging and skin cancer, such as damaging DNA, causing gene mutations, immunosuppression, oxidative stress, and inflammatory responses [5]. If the gene is mutated, it allows abnormalities in the DNA repair process, resulting in dysregulation of apoptosis and initiation of skin cancer. This situation is very worrying and can be a threat to humans, especially with global warming which increases the risk factors for skin cancer[4].

The sun exposure factor does not explain the cause of skin cancer that occurs on areas of skin that are not exposed to sunlight. This suggests that there are other factors that may contribute to the development of skin cancer, one of which is skin color. Apart from being caused by sun exposure, age, and skin color, skin cancer can also be affected by lifestyle, viruses, and the environment. Common symptoms experienced by people with skin cancer include painless protruding areas of skin that appear shiny with blood vessels on top, hard lumps with scaly tops; moles that change in size, color, irregular edges, more than one color, itching or bleeding; dark lesions on the palms of the hands, soles of the feet, fingertips or toes; or flat, flesh-colored or brown lesions like scars. The appearance of these symptoms will then be diagnosed by a dermatologist by following the mnemonic "ABCDE" which stands for Asymmetry, Borders, Color, Diameter, and Evolution [6].

Indonesian health profile reported that expert medical numbers are unevenly distributed, particularly in rural areas with substandard healthcare facilities. This is inversely proportional to the number of skin cancer cases especially in countries with high levels of sun exposure. Various procedures for identifying skin cancer that are non-invasive have been proposed, one of which is computer-based technology. The rapid development of technology, one of which is the use of artificial intelligence (AI) systems, can be an option in dealing with diagnostic process problems. In addition to being a solution as a non-invasive procedure, the implementation of artificial intelligence systems in the diagnosis process resources involved. The application of artificial intelligence technology may prove to be a significant breakthrough, potentially also assisting medical personnel in providing a second opinion.

Based on machine learning, the skin cancer classification system was designed using the Support Vector Machine (SVM) algorithm [7]. The seleccan also overcome the problem of the number of expert medical personnel to make a diagnosis [8]. With technology based on artificial intelligence systems, the effectiveness in diagnosing skin cancer can be improved by saving diagnosis time and human.

tion of this algorithm is based on the algorithm's reliable performance in the classification system. In research conducted by Faruk (2020), skin cancer classification was carried out using the Support Vector Machine (SVM). The research obtained an accuracy of 69.85% for the Support Vector Machine (SVM) algorithm [2]. Another research conducted by Monica (2020), the classification was also carried out using the Support Vector Machine (SVM) algorithm. The research used the dull razor method in the preprocessing step by removing noise in the form of hair on the patient skin's dermoscopy images which obtained an accuracy 96.25% [9]. In this research, the development will be done by using fractal method and Gray Level Co-Occurrence Matrix (GLCM) with homogeneity, energy, correlation features for feature extraction. Contrast enhancement and artifact removal were used to enhance image quality and reduce noise, specifically the appearance of hair on dermoscopic images [10]. The program can later be used for dermoscopy image identification of skin cancer with 2 classes, non-melanoma and melanoma [11].

2 Proposed Methodology

2.1 Materials and Methods

The system is designed using python programming language where used to identify medical record data of skin cancer dermoscopy images. Secondary datasets were collected and sorted by the International Skin Imaging Collaboration (ISIC), which consists of 9 classes. However, only two classes will be used in this study: Melanoma and Non-Melanoma, which is the accumulation of data from the Basal Cell Carcinoma (BCC) and Squamous Cell Carcinoma (SCC) classes. The data obtained will then be used in the training and testing process. The methods for the identification program will be explained below.

Preprocessing. Digital image processing is a technique used to enhance the quality of an image, facilitating its interpretation by humans or computer algorithms. It can be applied to both still images and moving images [12]. Each image object possesses a discernible difference in value that can be calculated mathematically, thus indicating distinctive characteristics that differentiate one object from another. The distinguishing characteristics of each object can be determined by analyzing its color, texture, or shape [13].

Resizing. Resizing is a standard method for image processing, whereby the dimensions of an image are uniformly altered to conform to a specified target size, without any portion of the image being removed. This process is more efficient and can only be achieved by considering the content of the image, rather than solely relying on geometric limitations, in Eq. (1) shows the formula of final resolution [10].

$$\text{Final resolution} = \frac{1}{4} \times \text{initial resolution} \tag{1}$$

Contrast Enhancement (CE). The fundamental principle of Contrast Enhancement (CE) is the modification of image information in order to optimize its suitability for a specific application [14]. CLAHE is a histogram-based method for enhancing contrast that expands the image contrast by distributing the gray-scale levels more evenly and limiting histogram amplification [15]. The objective of this process is to obtain a more contrast-rich image without compromising its quality [16].

Artifact removal. The dermoscopy images system offers a range of pixel sizes, intensities, and potential for the introduction of unwanted noise artifacts in the form of body hair. Such artifacts can impede the segmentation process in certain images and diminish the overall performance of the model. The process of removing the artifacts in question begins with the conversion of the dermoscopic image input, which is in RGB format, into a grayscale image. This is followed by a black hat transformation. The black hat transformation is a morphological transformation operation on binary images that accentuates dark objects against a bright background [17]. The black hat transformation yields an output image with darker elements than the surrounding pixels, thereby creating an artifact resembling hair. The resulting black hat transformation image is then processed through inpainting, whereby the threshold image from the black hat transformation and the original image are combined, and the hair artifact or structure is replaced with neighboring pixels [18].

Grayscaling. The process of grayscaling is defined as the conversion of a Red, Green, Blue (RGB) image, which has three channels, into a monochromatic image with a single channel. The objective of grayscaling is to achieve an appropriate contrast in the image. Contrast modification is achieved through grayscaling, which is used to reveal the characteristics of the image, facilitating the extraction of features using fractal methods, as shown in Eq. (2) [19].

$$S = 0,30(R) + 0,59(G) + 0,11(B) \tag{2}$$

Thresholding. The determination of the threshold or threshold value is of great consequence in the process of segmentation [20]. Thresholding is a widely used technique in image segmentation, whereby objects are separated from their backgrounds. Threshold, also known as an ambient, represents a value that has two regions on either side, which are below the threshold and above the threshold. The fundamental principle of thresholding is the grouping of pixels into binary categories based on the specified threshold value, thereby representing the object and background regions [21].

Feature extraction. The subsequent stage is the extraction feature, which is the stage of obtaining distinctive characteristics from the image. The selected extraction feature combines the results of the fractal feature extraction and the Gray-Level Co-occurrence Matrix (GLCM) as input for the classification process.

Fractal method. The term "fraktal" is defined as a fractal, which is a pattern that is self-similar and has a repeating structure. It can be described as either broken or irregular. The term fraktal was first introduced by Benoit B. Mandelbrot in 1973. An object that is fractal in nature exhibits a resemblance to itself at varying scales. This implies that the smaller components of the object will mirror the overall structure when observed as a whole [22]. The fractal dimension can be employed to quantify the roughness or smoothness of a curve within a given space, such as the border of skin cancer [23]. One

prevalent method for calculating the fractal dimension of an image is the Box Counting (BC) method, which focuses on calculating the relationship between the number of squares and the ratio of squares, which is expressed in the following equation.

$$N(R) = \frac{1}{R^D} \tag{3}$$

The following definitions are provided for clarification in Eq. (3). Where N(R) = number of squares that contain the image, and R = scale, and D = dimension.

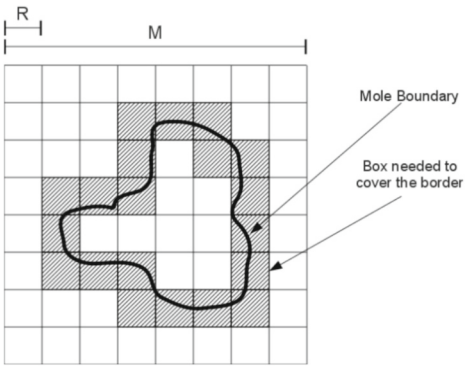


Fig. 1 Box counting method

Gray-Level Co-occurrence Matrix (GLCM), as shown in Fig. 1. Statistical method for forming features or characteristics that are not based solely on pixel values, but also on the relationships between neighboring pixels. In the implementation of the GLCM method, the GLCM value is defined as a matrix that expresses the spatial distribution between two neighboring pixels with intensity values i and j, a distance d between them, and an angle θ between them, which is represented by the notation Pd, θ (i, j). The angle between the two neighboring pixels can be classified into eight distinct directions: 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315° (Widyaningsih, 2016). The Gray Level Co-occurrence Matrix (GLCM) values obtained are then used to calculate the unique elements of the relevant objects, employing the contrast, correlation, homogeneity, and energy metrics.

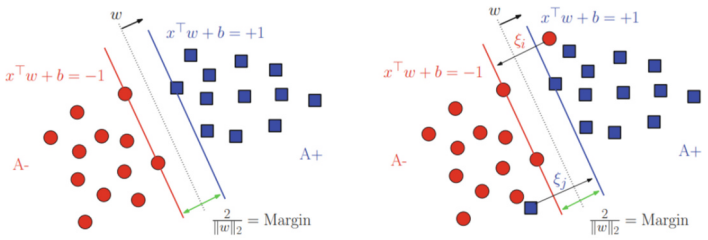
Resampling Data. In the course of data processing, it is not uncommon to encounter an imbalanced dataset with respect to each class. An imbalanced data set is one in which the number of instances within a given class is significantly higher (majority class) or lower (minority class) than that of other classes. The presence of imbalanced data presents a significant challenge for data processing and analysis tasks such as classification, clustering, prediction, and others, as the system is not designed to accommodate class distribution considerations. The use of resampling involves modifying the number of data points in each class until a balanced distribution is achieved across all classes [24]. The SMOTEENN method represents a sophisticated combination of undersampling and oversampling techniques, simultaneously applied to the data set. The underlying

principle of this method is the integration of two distinct approaches: firstly, the ability of SMOTE to generate synthetic examples within minority classes, and secondly, the capacity of ENN to remove observations from both classes that are identified as having different classes from their nearest majority neighbors [25].

Cross Validation. K-fold cross-validation is a data partitioning strategy that allows for the development of a model with robust performance on new data. This stage is also conducted to enable generalization of the data, whereby the performance of the model is evaluated using data that has not been employed during the training process. The dataset is divided into several subsets for training and testing purposes. The most effective approach for this is to use k-fold cross-validation. The k-fold cross-validation method was selected as a means of overcoming the limitations of the sample data. The k-fold cross-validation method is employed to guarantee that each data point from the data set is represented in both the training and testing sets.

Modeling. Support Vector Machines (SVMs) represent a prominent approach within the domain of machine learning, particularly suited to the task of data classification. The method was first introduced by Vapnik in 1992 [26]. The SVM method operates on the principle of structural risk minimization (RSM), which aims to identify the optimal hyperplane (line) for separating two classes in the input space [27]. The SVM operates as a linear classifier, separating two data classes with a linear line. However, in many cases, the two classes cannot be divided linearly, prompting the development of the SVM with the addition of the kernel trick, which elevates the space to a higher dimension [26].

A hyperplane is a separating line that optimally distinguishes between two classes. The dimension of the hyperplane is d-1 in a d-dimensional vector space. Consequently, a two-dimensional hyperplane in a two-dimensional space is a one-dimensional line, while a three-dimensional hyperplane in a three-dimensional space is a plane. The optimal hyperplane is identified by locating the margin value and the maximum point. The margin represents the distance between the hyperplane and the nearest pattern from each class. The pattern that is closest to the hyperplane is referred to as the support vector [26].



(a) linearly separable (b) non-linearly separable

Fig. 2 Model kernel Support Vector Machine (SVM)

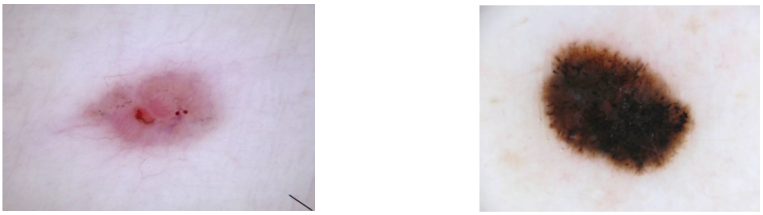
In Fig. 2. provides an illustration of data that cannot be divided in a linear or non-linear separable manner. Subsequently, SVM was further developed by incorporating the kernel trick, thereby elevating the space to a higher dimension [26]. The kernel trick is a crucial component of SVM, as it addresses the need to handle data that cannot be separated linearly in practice. A kernel is a function that maps data to a higher-dimensional vector space, which its formula shown in Table 1. The most commonly employed kernels are polynomial, Gaussian (RBF), linear, and sigmoid, each of which possesses a distinct mathematical function [28].

Table 1 Kernel Tricks Function

Jenis Kernel	Definisi
Polinomial	$K(\vec{x}_i, \vec{x}_j) = (\vec{x}_i . \vec{x}_j + c)^p$
RBF	$K(\vec{x}_i, \vec{x}_j) = \exp(-\sigma \lvert\lvert \vec{x}_i - \vec{x}_j \rvert\rvert^2)$
Linear	$K(\vec{x}_i, \vec{x}_j) = \vec{x}_i . \vec{x}_j$
Sigmoid	$K(\vec{x}_i, \vec{x}_j) = \tanh(\alpha \vec{x}_i . \vec{x}_j + \beta)$

2.2 Data Input

This study will only use 2 classes: melanoma and non-melanoma, as shown in Fig. 3, which are accumulated data from the Basal Cell Carcinoma (BCC) and Squamous Cell Carcinoma (SCC) classes from the dataset. The total image data to be used is 589 in the non-melanoma class and 454 in the melanoma class from 2357 images in the initial dataset. The set of images in the dataset also has different resolutions, which range from 600×450 to 3072×2304. The dataset will then be input to the identification program basic design and identification program development.



(a) (b)

Fig. 3 Examples of input data, (a) non-melanoma, filename ISIC_0024436, (b) melanoma, filename ISIC_0000013

3 Results and Discussion

The proposed methodology for identification program development is shown using block diagrams Fig. 4.

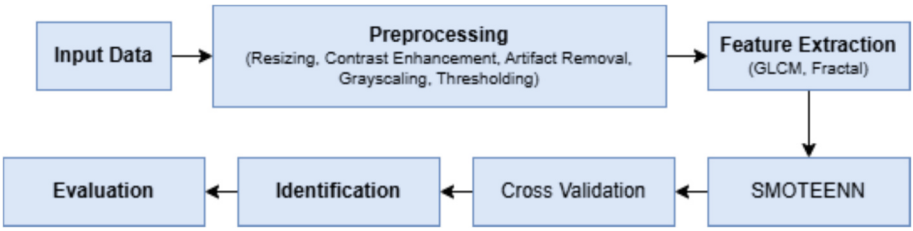


Fig. 4 Block Diagram of Identification Program Development

Preprocessing. The preprocessing stage has 5 processes, which include resizing, contrast enhancement, artifact removal, grayscaling, and thresholding. The first unique image processing added in the preprocessing stage is Contrast Enhancement (CE). This stage is aimed at enhancing features in images that have low contrast. Through this stage, the problem of contrast differences between images in the dataset can also be resolved. Prior to Contrast Enhancement (CE), a resizing process will be carried out to reduce the computational load, so that the model training process can run efficiently and optimally. Through this resizing process, the image will be resized to produce an output resolution of 0.25 or $\frac{1}{4}$ times the original resolution, as shown in Fig. 5.

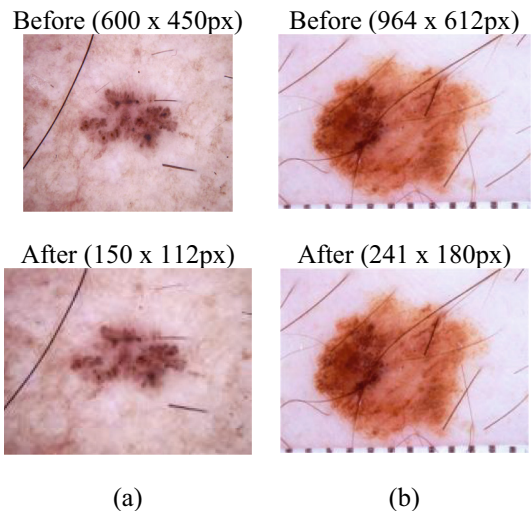


Fig. 5 Resizing process on dataset, (a) non-melanoma, filename ISIC_0024403, (b) melanoma, filename ISIC_0000171

The results of the resizing process will be input to the Contrast Enhancement (CE) process. At this stage, the Contrast Enhancement (CE) process is performed using Contrast Limited Adaptive Histogram Equalization (CLAHE). CLAHE is a variant of Adaptive histogram equalization (AHE) that handles excessive contrast amplification and can operate on small areas of the image i.e. tiles, rather than the entire image. Adjacent tiles are then merged using bilinear interpolation to remove artificial boundaries. In the Contrast Enhancement (CE) stage, the image will be converted into an array first so that it can be processed using CLAHE through the numpy module. The next step is to change the channel order in the image from Red, Green, Blue (RGB) to Blue, Green, Red (BGR), as shown in Fig. 6.

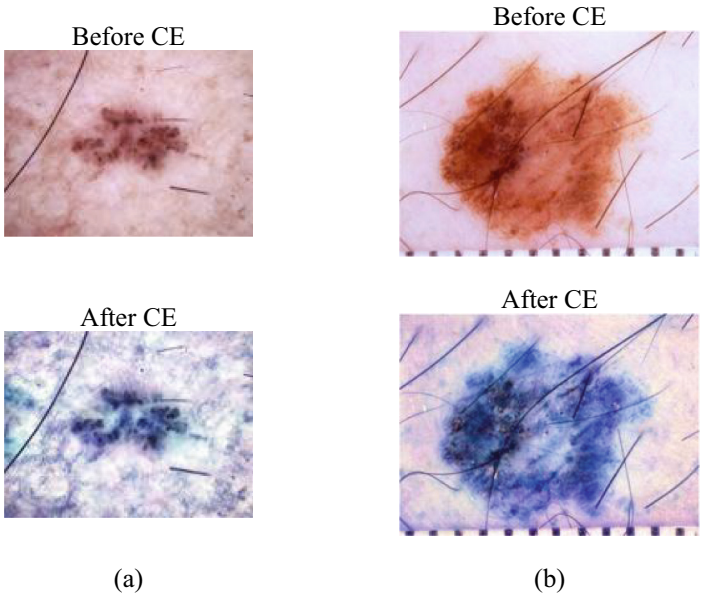


Fig. 6 Contrast Enhancement process on dataset, (a) non-melanoma, filename ISIC_0024403, (b) melanoma, filename ISIC_0000171

In the image Fig. 6, it can be seen that the image becomes blue as a result of the process of changing the channel order to Blue, Green, Red (BGR). Through the stages of channel reordering and the implementation of Contrast Limited Adaptive Histogram Equalization (CLAHE), a sharper output image can be produced and can show unique elements in the image. The combination of the channel reordering process and the implementation of Contrast Limited Adaptive Histogram Equalization (CLAHE) can provide optimal contrast improvement in the image because it can simultaneously sharpen the image and display unique elements in the image.

The next image processing added to the preprocessing stage for further development of the identification program is artifact removal. The use of this stage is aimed at reducing noise in the image in the form of hair. This is because artifacts such as hair can be a disturbing factor in the segmentation results of some images and can reduce the overall performance of the model.

The steps to be taken for the implementation of this artifact removal process begin with converting the original BGR image which is the output of the Contrast Enhancement (CE) process into a grayscale image. Followed by the use of blackhat transformation which is intended to emphasize dark objects against a light background. This blackhat transformation can produce an output image containing elements that are darker than the surrounding pixel values, so that the transformation results are obtained to obtain hair-like artifacts. The next step after the blackhat transformation process is the thresholding process. Through this process, a binary output image is produced with black color as the background and white color on the artifact, namely hair feathers. The threshold results will be inpainted or recolored by replacing the artifact or hair structure with neighboring pixels.

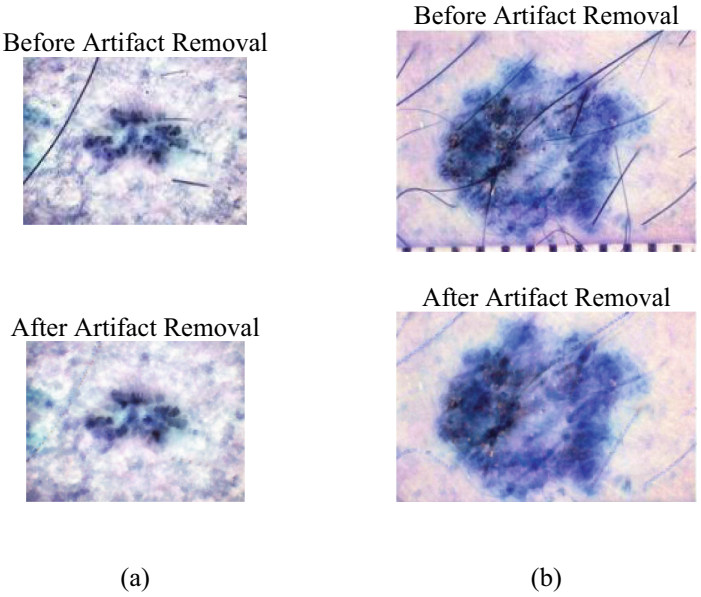


Fig. 7 Artifact removal process on dataset, (a) non-melanoma, filename ISIC_0024403, (b) melanoma, filename ISIC_0000171

In the image Fig. 7, it can be seen that the image at the top initially contained artifacts in the form of hair. After the image inpainting process, the output image is produced with the artifacts that have been camouflaged. The artifact removal method is very useful in the implementation of image processing that has many disturbing factors in the form of hair. This method can eliminate hair, resulting in an image that is minimal or even clean from hair interference. The use of this method can provide optimal output results

because image processing does not involve disturbing factors in the form of hair that can affect system performance.

Another image processing added to the preprocessing stage is grayscaling. This stage is intended to convert a Red, Green, Blue (RGB) image that has 3 channels into a gray image that only has 1 channel. This process is done by multiplying the R matrix, G matrix, and B matrix then summing the results into a new matrix. In addition, the use of grayscaling is done so that it can be an input to the feature extraction that will be added, namely Gray Level Co-Occurrence Matrix (GLCM). In this research, the coefficients on the Red, Green, Blue channels were used, namely 0.30, 0.59, and 0.11, respectively, as shown in Fig. 8.

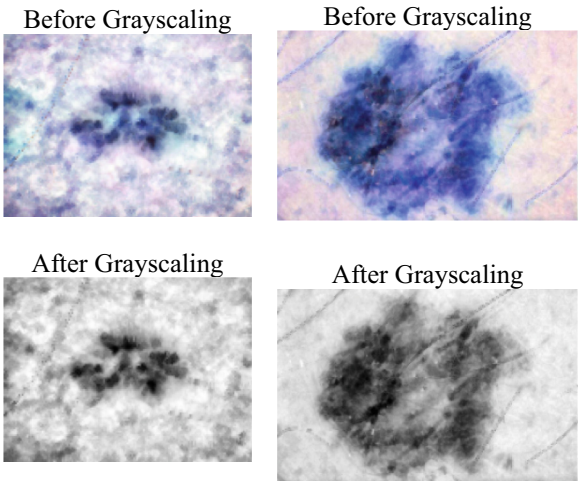


Fig. 8 Grayscaling process on the dataset, (a) non-melanoma, filename ISIC_0024403, (b) melanoma, filename ISIC_0000171

The thresholding stage is performed to separate the main object pixels from the background pixels based on the pixel intensity value. The output image from this thresholding process is a binary image that will be used as input to the fractal method as an extraction feature.

The working principle of the function is that if the pixel intensity is greater than the predefined threshold value, then the pixel value is fixed. However, if the pixel intensity is lower than the predefined threshold value, then the pixel value is converted to the lowest pixel value of 0 which is black, as shown in Fig. 9.

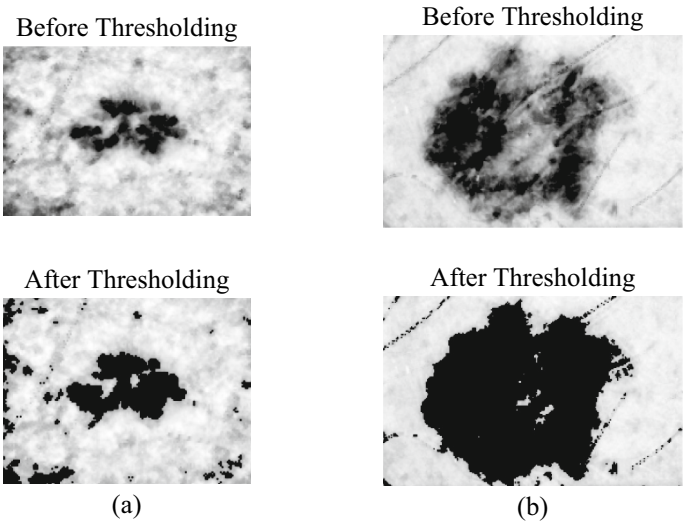


Fig. 9 Thresholding process on dataset, (a) non-melanoma, filename ISIC_0024403, (b) melanoma, filename ISIC_0000171

The non-melanoma image uses 160 as the threshold value parameter, so the image pixel values above 160 do not change. While the image pixel values that are below the threshold value are converted to the lowest pixel value of 0 which is black. Through this process, a gray output image is produced in the part with pixel values above the threshold value and a black part in the pixel values below the threshold value. The process can immediately mark the skin area that is skin cancer and separate it from the background, where the skin cancer area is marked with a black area. While the background is marked with a grayish colored area. The same process occurs for melanoma images that use 180 as the threshold value parameter.

Feature extraction. The next stage after the preprocessing stage is feature extraction which are fractal method and Gray-Level Co-occurrence Matrix (GLCM) with features in GLCM such as homogeneity, energy, and correlation.

The use of the fractal method with the specification of implementing the box counting method is intended to estimate the fractal dimension of an object by measuring the number of boxes needed to cover it at different scales or resolutions. The fractal method used shows unique characteristics in each class that has a threshold value below 0.1. The selection of the threshold value is based on the thresholding process in the preprocessing stage. This stage provides an output image that shows specifically the skin cancer part in the black area, in accordance with the selection of the threshold value of 0.1 which is close to the black color. However, the results of feature extraction on both data frames still have a range difference that is not much different from one class to another. Therefore, additional features were extracted, namely Gray-Level Co-occurrence Matrix (GLCM) with features in GLCM such as homogeneity, energy, and correlation. The selection of these features is intended as additional information for the system so that it is expected to improve system performance.

Gray-Level Co-occurrence Matrix (GLCM) was selected to provide additional information for the system to learn due to its ability to recognize patterns and spatial distribution in different images. GLCM can represent the relationship between two neighboring pixels that have grayscale intensity, distance and angle. GLCM features such as homogeneity, energy, and correlation will be used. The selection of these features is based on their respective uses, homogeneity to measure the homogeneity of a pixel point with the pixel values around that point. The use of energy to determine the uniformity between one pixel and another. Energy will be high if the pixel value is similar to other pixels, while the energy value will be low if the pixel value is not similar or heterogeneous. Finally, the use of correlation aims to measure the linearity of a number of pixels.

Data resampling. As mentioned, data resampling was performed using the SMOTEENN function. The following is the amount of data after the implementation of the SMOTEENN function on each class shows in Table 2.

Table 2 Results of the Implementation of the SMOTEENN Function for Non-melanoma and Melanoma Classes in the Program

Class	Before	After
Non-melanoma	589	407
Melanoma	454	435

Cross validation. After resampling data, the k-fold cross validation stage is used as one of the dataset splitting methods that divides the dataset into 2 parts, namely training data and test data before training and testing on the algorithm to be built. The identification algorithm that will be used in identification program development is Support Vector Machine (SVM) on linear, RBF, polynomial, and sigmoid kernels that show in Fig. 10.

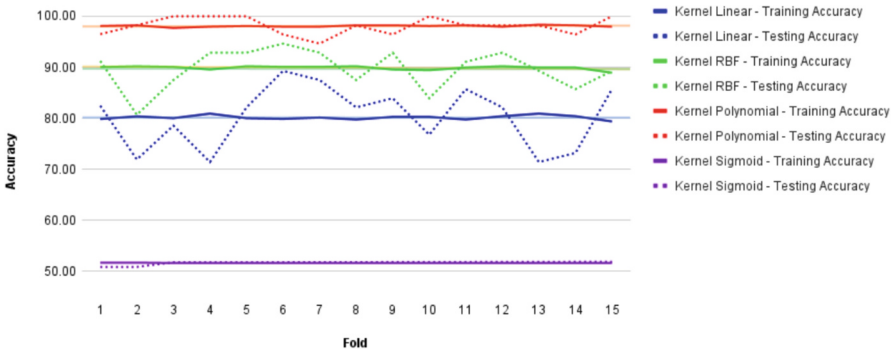


Fig. 10 Support Vector Machine (SVM) Algorithm Accuracy Values using K-fold cross-validation

Identification. The next stage after k-fold cross validation as a data partitioning strategy is to retrain the system using the data set that has the best accuracy. The selection of data sets with the best accuracy is based on testing accuracy which has the highest accuracy compared to other folds.

Evaluation. The retrain results will be evaluated using accuracy, precision, recall, and f1-score values. The use of advanced evaluation methods namely precision, recall, and f1-score is intended as additional validation of the evaluation results based on accuracy values. This is because accuracy can provide a high output value even though the model used has poor performance. Therefore, it is necessary to calculate precision, recall, and f1-score to see if the resulting accuracy is in accordance with the system performance. The value of accuracy, precision, recall, and f1-score is shown on the Table 3.

Table 3 Evaluation Table of Support Vector Machine (SVM) Algorithm in Linear, RBF, Polynomial, and Sigmoid kernel

Kernel	Evaluation				
	Training Accuracy (%)	Training Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Linear	79.90	89.29	89.46	89.29	89.26
RBF	90.08	94.64	94.71	94.64	94.64
Polynomial	97.71	100.00	100.00	100.00	100.00
Sigmoid	51.65	51.79	26.82	51.79	35.34

Through the table, it can be seen that the kernel that has the best accuracy is the polynomial kernel, followed by the RBF, linear, and sigmoid kernels. The polynomial kernel in both programs provides optimal results. This is because the polynomial kernel can adjust to the data pattern of the feature extraction results. In the basic design program, the accuracy results are still quite low, namely 64.38% in training results and 74.24% in testing results. Meanwhile, the identification program development reached 97.71% in training results and 100% in testing results with the use of polynomial kernels in the Support Vector Machine (SVM) algorithm. This proves that the development program has a more optimal performance, as shown in Fig. 11.

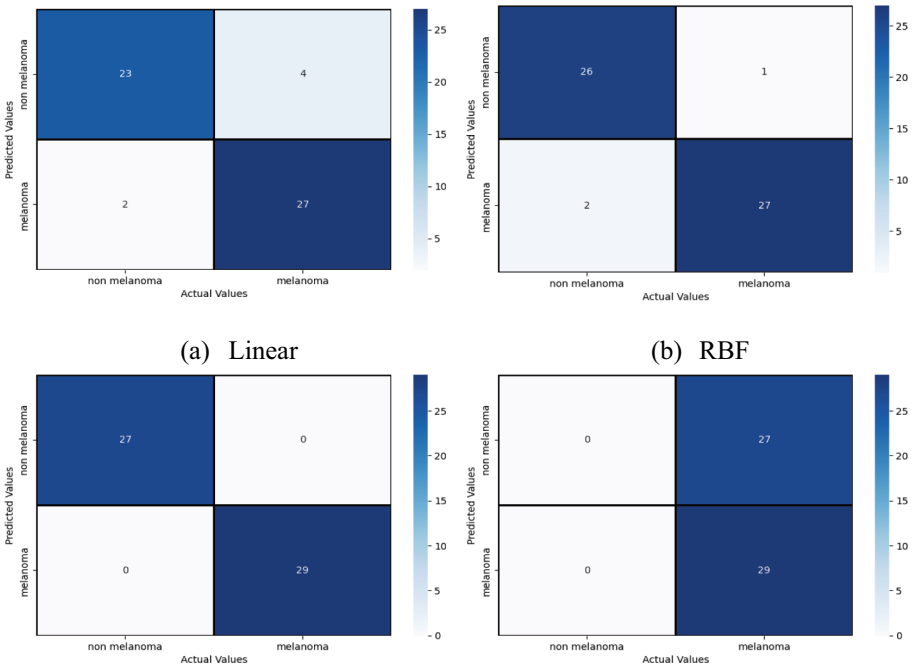


Fig. 11 Confusion matrix of Identification Program Development Using Support Vector Machine (SVM) Algorithm, (a) Linear Kernel, (b) RBF Kernel, (c) Polynomial Kernel, (d) Sigmoid Kernel

4 Conclusion

Based on the results of the research, it can be concluded that the performance of the system with the use of fractal method and additional feature extraction can provide further information for the system to be able to optimize its performance. In the form of additional preprocessing and feature extraction stages until the accuracy results on the same kernel are 97.71% in the training process and 100% in the testing process. The accuracy obtained from the kernel polynomial represents the optimal result for the dataset and image processing techniques employed. Further development of this research could be achieved through integration with a Graphical User Interface (GUI) system, thus enhancing its utility for supporting medical operations.

References

1. Slominski A, Wortsman J (2000) Neuroendocrinology of the skin1. Endocr Rev 21:457–487. https://doi.org/10.1210/edrv.21.5.0410
2. Faruk M, Nafiiyah N (2020) Klasifikasi Kanker Kulit Berdasarkan Fitur Tekstur, Fitur Warna Citra Menggunakan SVM dan KNN. Telematika. 13:100–109
3. Apalla Z, Lallas A, Sotiriou E, Lazaridou E, Ioannides D (2017) Epidemiological trends in skin cancer. Dermatol. Pract. Concept. 7. https://doi.org/10.5826/dpc.0702a01

4. Narayanan DL, Saladi RN, Fox JL (2010) Review: Ultraviolet radiation and skin cancer. Int J Dermatol 49:978–986. https://doi.org/10.1111/j.1365-4632.2010.04474.x
5. Queen L (2017) Skin cancer: causes, prevention, and treatment
6. Chatterjee S, Dey D, Munshi S (2015) Mathematical morphology aided shape, texture and colour feature extraction from skin lesion for identification of malignant melanoma. In: 2015 international conference on condition assessment techniques in electrical systems (CATCON), pp, 200–203. IEEE. https://doi.org/10.1109/CATCON.2015.7449534
7. Ghaderi S (2023) Fractal dimension image processing for feature extraction and morphological analysis: Gd3+/13X/DOX/FA MRI nanocomposite. J Nanomater 2023:8564161
8. Jaleel JA, Salim S, Aswin R (2012) Artificial neural network based detection of skin cancer. Int J Adv Res Electr Electron Instrum Eng 1
9. Monika MK, Vignesh NA, Kumari CU, Kumar M, Lydia EL (2020) Skin cancer detection and classification using machine learning. Mater Today Proc 33:4266–4270
10. Avidan S, Shamir A (2023) Seam carving for content-aware image resizing. In: Seminal graphics papers: pushing the boundaries, vol 2, pp 609–617
11. Nikitin V, Danilov V (2023) Integration of fractal dimension in vision transformer for skin cancer classification. Electron Control Syst 2:15–20
12. Effendi M, Fitriyah F, Effendi U (2017) Identifikasi Jenis dan Mutu Teh Menggunakan Pengolahan Citra Digital dengan Metode Jaringan Syaraf Tiruan. J Teknotan 11:67
13. Widyaningsih M (2016) Identifikasi kematangan buah apel dengan gray level co-occurrence matrix (GLCM). J Saintekom Sains, Teknol Komput dan Manaj 6:71–88
14. Qi Y, Yang Z, Sun W, Lou M, Lian J, Zhao W, Deng X, Ma Y (2021) A comprehensive overview of image enhancement techniques. Arch Comput Methods Eng pp 1–25
15. Adzkia M, Arland F, Setiawan AW (2022) Deteksi Pneumonia Menggunakan Citra Sinar-X Paru berbasis Residual network. J Teknol Inf dan Ilmu Komput 9:373–380
16. Hayati M, Muchtar K, Maulina N, Syamsuddin I, Elwirehardja GN, Pardamean B (2023) Impact of CLAHE-based image enhancement for diabetic retinopathy classification through deep learning. Procedia Comput Sci 216:57–66
17. Wang G, Wang Y, Li H, Chen X, Lu H, Ma Y, Peng C, Wang Y, Tang L (2014) Morphological background detection and illumination normalization of text image with poor lighting. PLoS ONE 9:e110991
18. Basak H, Kundu R, Sarkar R (2022) MFSNet: a multi focus segmentation network for skin lesion segmentation. Pattern Recognit 128:108673
19. Akbar IH, Sassongko PS (2021) Pengenalan Pneumonia Pada Anak Menggunakan Jaringan Syaraf Tiruan Extreme Learning Machine. In: Seri Prosiding Seminar Nasional Dinamika Informatika
20. Sharma SR, Alshathri S, Singh B, Kaur M, Mostafa RR, El-Shafai W (2023) Hybrid multilevel thresholding image segmentation approach for brain MRI. Diagnostics 13:925
21. Rani G, Khandelwal S, Kundu N, Dhaka VS (2023) Automatic segmentation and classification of skin cancer cells using thresholding and deep learning based techniques. In: 2023 international conference on emerging trends in networks and computer communications (ETNCC), pp 1–6. IEEE
22. Putra D (2010) Pengolahan citra digital. Penerbit Andi
23. Carbonetto SH, Lew SE (2010) Characterization of border structure using fractal dimension in melanomas. In: 2010 Annual international conference of the IEEE engineering in medicine and biology. IEEE, pp 4088–4091
24. Indrawati A (2021) Penerapan Teknik Kombinasi Oversampling Dan Undersampling Untuk Mengatasi Permasalahan Imbalanced Dataset. JIKO (Jurnal Inform. dan Komputer). 4:38–43
25. Vairetti C, Assadi JL, Maldonado S (2024) Efficient hybrid oversampling and intelligent undersampling for imbalanced big data classification. Expert Syst Appl 246:123149

26. Nugroho AS, Witarto AB, Handoko D (2003) Support vector machine. Proceeding Indones. Sci. Meeiting Cent. Japan
27. Religia Y, Rusdi A, Romli I, Mazid A (2019) Feature extraction Untuk Klasifikasi Pengenalan Wajah menggunakan support vector machine Dan K-Nearest Neighbor. Pelita Teknol 14:85–92
28. Lee Y-J, Yeh Y-R, Pao H-K (2011) Introduction to support vector machines and their applications in bankruptcy prognosis. In: Handbook of computational. Springer, finance, pp 731–761 (2011)

Shannon Entropy as Feature for Improved Prediction of Cardiac Risk Events in ECG Signals

Giovanny Andrés Piedrahita Solórzano[1](✉), José Luis Ramírez Arias[2], and Daniel Botero Rosas[3]

[1] Institución Universitaria Politécnico Grancolombiano, Bogota, Colombia
gapiedrahita@poligran.edu.co
[2] Universidad Militar Nueva Granada, Cajicá, Colombia
[3] Universidad de la Sabana, Bogotá, Colombia

Abstract. About 50% of all cardiovascular-related deaths are the result of a sudden cardiac death (SCD) event, which is the abrupt loss of consciousness within less than one hour after the onset of symptoms. While the SCD is the ultimate event, it is triggered by other cardiac events characterized by alterations in the normal rhythm of the heart. This research explores a novel approach for analyzing the electrocardiogram (ECG) signals and machine learning algorithms to predict events such as ventricular fibrillation and generate alerts early enough to allow access to healthcare. This approach extracts the Shannon entropy as a time-variable feature from the ECG samples organized on multiple time-windows, enhancing SCD prediction in terms of lead time and accuracy respect to current results reported in literature. The experiments in this study were made with ECG records from the Physionet public ECG databases, taking the samples and annotations to train and evaluate several supervised learning algorithms, such as binary decision trees, convolutional neural networks (CNN), and support vector machines (SVM). Results show that support vector machines (SVM) have the best performance to predict the SCD event, with high performance metrics evaluated 15 min before the onset. Results show that calculating Shannon entropy from raw ECG signals has a better performance than other time-domain features extracted from the ECG, and from heart rate variability (HRV) features, allowing for prediction up to 15 min before the SCD event happens reaching 92.5% accuracy. The proposed model, trained with the SVM algorithm, demonstrates a competitive and improved classification accuracy and F1 score (95%), measured for 5 min before the onset to compare it with existing methods (92.31%). While most of the referenced papers report results in a time span of 5 min, this research contributes to the advancement of automatic SCD prediction techniques extending the time up to 15 min with a good prediction, detailing the method to replicate it with longer times. As the model works with one feature to represent the ECG signal, the computational requirements will be lower for potential applications in wearable monitoring devices.

Keywords: Sudden cardiac death · automatic detection · Shannon entropy

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 219–234, 2026.
https://doi.org/10.1007/978-981-95-9233-3_17

1 Introduction

Sudden cardiac death (SCD) is the cause of about half of deaths related to cardiovascular diseases, being the primary cause of mortality globally. Sudden cardiac death (SCD) is defined as unexpected death due to cardiac causes, occurring within 1 h from the onset of symptoms in individuals with no known prior conditions that would limit their lifespan [1–3]. Despite the decrease in mortality rates from coronary artery disease, there are still high death rates from SCD [4, 5]. In the United States, between 250,000 and 300,000 people die suddenly every year, and the incidence of SCD in North America and Europe ranges between 50 and 100 SCD per 100,000 inhabitants [5–8]. It is common for many of these deaths to occur because the person is not aware of the existence of a cardiac pathology or, if they know about it, perform actions that can overload the system and therefore cause death [9]. The most frequent consequence of cardiac abnormalities that lead to an SCD event is ventricular fibrillation, therefore, it is challenging to set up effective mechanisms for identifying possible events that may trigger sudden death or to establish mechanisms for their prevention in the general population.

Scientific community has been interested in understanding the reasons for an SCD event to happen, and the research has been migrating over the years from a biological perspective to the field of automatic prediction, given the diversity of techniques available and access to computing power for implementing them. The results in predicting the events before SCD have been improving by applying machine learning models and algorithms to data gathered from patients who have had tachycardia, arrhythmia, fibrillation, or another abnormal heart rhythm. Features in time and frequency domains have been used as features, as well as nonlinear and morphological analysis to train the algorithms, producing an extensive literature reviewed by representative authors in the field [10–14], Table 1 summarizes some of the results relevant to this research as they highlight prediction techniques, reference public ECG signal databases and use standard classification metrics.

From the literature, time domain analysis has been one of the most used for data preprocessing and feature extraction, because it's directly related to the way in which experts examine the ECG signal to diagnose the pathologies. Measuring the interval from beat to beat (R-R intervals) leads to a heart variability rate time-series, and features such as mean and standard deviation are used to train the algorithms [22, 23]. Other analysis techniques and domains have been proposed, where the features selected include the ratios of average power in frequency bands (frequency domain) [24], the entropy as a measure of the complexity of the signal (nonlinear techniques) [25], the average power and entropies in bispectrum analysis [13], or a composition from all of them. In [19] a set of 50 features from four different domains is the input for the classification algorithm, and shows the variation in prediction accuracy depending on the number of features. While working with an extended set of features can increase the prediction confidence, its hardware implementation will require more powerful processors and memory, limiting to on-site devices like monitors for intensive care units.

These works follow a process shown in Fig. 1, using the selected features to train the algorithm, and later, to make predictions extracting the same features from incoming new ECG data.

Table 1. Literature results for automatic SCD prediction.

Year	Title	Results	Database
2016	Sudden cardiac death (SCD) prediction based on nonlinear heart rate variability features and SCD index [15]	Four minutes earlier Accuracy of 97.3%	MITBIH Database
2018	A Novel Wavelet Transform-Homogeneity Model for Sudden Cardiac [16]	20 min prior to the onset with a high accuracy of 95.8%	Normal Sinus Rhythm (NSR) SCDH database
2018	A time local subset feature selection for prediction of sudden cardiac death from ECG signal [17]	SCD can be predicted 12 min before the onset; accuracy of 83%	Sudden Cardiac Death Holter Normal Sinus Rhythm
2019	An optimal strategy for prediction of SCD [18]	Sensitivity, specficity, and accuracy of 84.24%, 85.71%, and 82.85%, respectively	MITBIH Database
2020	Heart rate variability feature selection method for automated prediction of sudden cardiac death [19]	Sensitivity and specficity of 88.9%, and 93,5%, respectively, 5min before the event	Medtronic 1.0 Database
2021	Detection of ventricular arrhythmia using hybrid time–frequency-based features and deep neural network [20]	Sensitivity and specificity of 98,8%, and 99,3%, respectively, 5 min before the event	CUDB and VFDB Databases
2023	Intelligent prediction of sudden cardiac death based on multi-domain feature fusion of heart rate variability signals [21]	Accuracy of 92,11% 10 min prior to the onset	Normal Sinus Rhythm (NSR) SCDH database

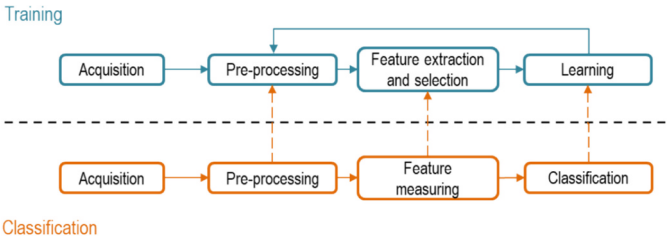


Fig. 1. Training and classification process for automatic prediction.

As Table 1 shows, as well as other SCD prediction reviews [13, 26], results are reported mostly for 1 and 5 min before the onset, reaching sensibility, specificity, and accuracy levels higher than 80% up to five minutes before the onset. Still, it is desirable to reach higher levels of confidence in predicting cardiac events over an extended period before the onset, and there is a lower number of results from working with other non-standard times, such as 10 or 13 min before the onset [17, 21]. Reaching an acceptable prediction more than ten or fifteen minutes could benefit potential patients by giving them more time to get opportune healthcare. On the other side, patients with Implantable Cardioverter Defibrillators (ICD) would have the chance of improving their quality of life; by decreasing the times they get an electrical stimulus because of a false positive detection.

This work aims to predict the onset up to 15 min before the onset, evaluating predictor's performance for each minute in that period, and propose the use of the Shannon entropy as unique feature to extract from the ECG signal. The process addresses two shortcomings identified from literature contributing in the following manner. (1) The advantages of using the Shannon entropy over other time-domain features to train the classification model. (2) The extraction of features from multiple time-windows to increase the performance of the classification model. (3) To achieve high performance over a longer time before the onset. (4) To evaluate performance of different classification algorithms for determining the best among them. (5) To decrease computational requirements for a hardware that makes real-time predictions. The performance of the proposed model is evaluated using accuracy and F1 score metrics, showing results that promise a more precise and earlier onset detection, increasing the reaction time for potential patients to access healthcare.

2 Method

This research is approached in three phases: dataset preparation, in which a detailed time-window management is proposed to generate the training datasets, algorithm experimentation, and evaluation. This evaluation extends the traditional reports on 1 and 5 min before the onset, to every minute from 1 to 15 min before the onset, to evaluate the decreasing performance as data is gathered farther from the onset. The following subsections describe with more detail the phases and how the result of each one produces inputs for the others.

2.1 Databases and Tools

Regarding data acquisition, there are several open datasets of EKG records available for research, most of them in Physionet [27], such as the spontaneous ventricular tachyarrhythmia, MIT-BIH normal sinus rhythm (https://www.physionet.org/content/nsrdb/1.0.0/), MIT-BIH Atrial Fibrillation (https://physionet.org/content/afdb/1.0.0/), and Sudden Cardiac Death Holter databases (https://physionet.org/content/sddb/1.0.0/) [28]. These databases include labels and annotations from experts that diagnose the patient's pathology, and records from healthy subjects, necessary for algorithm training in supervised learning techniques. This research has taken three databases to build datasets for training and testing the prediction algorithm, mainly because they included long-term records, easing the selection of onset points that provided a large enough number of samples before it to produce the dataset.

- MIT-BIH Atrial Fibrillation Database [27]: this database contains 10-h ambulatory ECG recordings from 25 subjects. All the recordings include atrial fibrillation and provide annotation files with beats and cardiac events.
- Sudden Cardiac Death Holter Database (SCDHD) [29]: this database contains 23 Holter recordings of 4 to 24 h length, from 18 subjects with sustained ventricular tachyarrhythmia, and 4 subjects with atrial fibrillation.
- MIT-BIH Normal Sinus Rhythm Database [30]: it includes 18 long-term ECG recordings from 5 men aged 26 to 45, and 13 women, aged 20 to 50.

To work with the databases, a custom Matlab application was created for invoking functions from the WFBD toolbox [31], extract information from samples and annotations, build the dataset, and finally, train the algorithms. Figure 2 shows the interface and main functions.

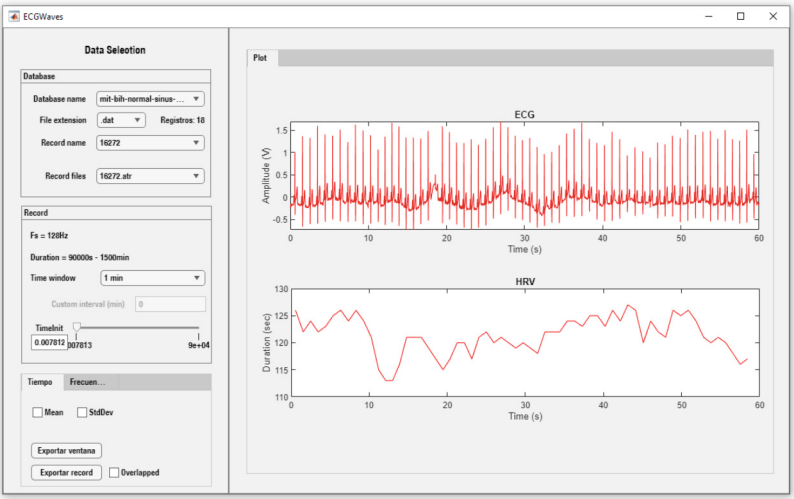


Fig. 2. Matlab application's UI.

The application was developed to ease the process of extracting the relevant information for this study, as the databases contain a significant amount information for each record. The main functions programmed in the application are the following:

- Selecting and loading a local Physionet database.
- Selecting an individual record and loading samples for ECG waveform visualization.
- Computing and visualization of HRV signal.
- Selecting an annotation file and extracting annotation types for filtering.
- Storing the information of a selected sample in a list of "onset samples".
- Extracting time-windows according to the method described in Sect. 2.2.
- Extracting features for dataset creation, storing them in separate data files.
- Classifier selection and training, storing images for performance metrics.

2.2 Time-Windows Management and Datasets Preparation

This work proposes a different approach for processing the features by training the model with multiple length time-windows. The following notation is defined in Fig. 3:

- t_{on}: onset time, measured from the record start.
- n: the total number of observations or samples labeled with an onset.
- k: length of the time-window to calculate the feature.
- i: minute before the onset, when a new observation is generated for the training dataset.

- j: minute before the onset, when the prediction is made. This eases the organization of testing datasets.
- f: number of features extracted for a time-window.
- In the training process, a dataset is constructed with the following procedure:
- For each t_j in an observation, the group of f features from different domains is calculated for each k-minute time-window. Therefore, each observation has k x f features.
- The feature extraction is replicated every minute from 1 to i-minutes before the onset, obtaining a dataset with n x i rows. The database must have enough samples during k minutes before the onset is detected.
- The dataset has now n x i rows and k x f columns. The observations selected from MIT-BIH Arrhythmia Database and Sudden Cardiac Death Holter Database are labeled with 1, while the ones selected from MIT-BIH Normal Sinus Rhythm Database are labeled with 0.

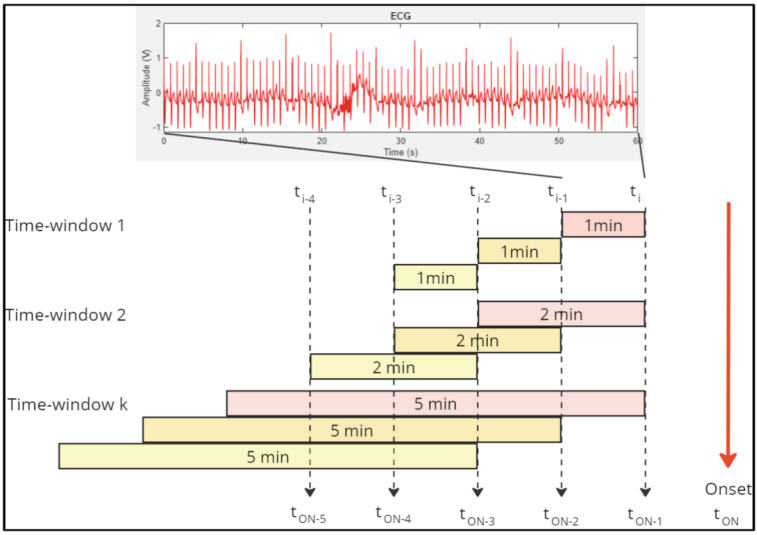


Fig. 3. Time-windows management

Applying this process for a maximum of $i = 15$, a vector of features is constructed for obtaining a dataset with nx15 observations (rows) and 5x3 features (columns), ensuring there is the same number of observations for each category or class in the training set. A total of 100 observations were selected as cardiac events from Physionet databases labeling them with 1, that were extended to 1500 by this process.

2.2.1 Shannon Entropy as Time-Domain Feature for Classification

Time domain features extracted from ECG and HRV signals are extensively used for implementing prediction algorithms. Alternatively, this research proposes the use of the Shannon entropy as a feature, calculating it on the multiple time-windows discussed previously. Table 2 shows the notation and signals used for obtaining the features from the ECG (*ecg(n)*) and HRV (*RR(n)*), signals. Mean and standard deviation are also extracted as references to compare the algorithm's performance to other features from time-domain.

Table 2. Features calculation.

Feature	Formula
Mean	$ECG_m = \frac{1}{N}\sum ecg(n)$
	$HRV_m = \frac{1}{N}\sum RR(n)$
Standard deviation	$ECG_{SD} = \sqrt{\frac{1}{N}\sum(ecg(n) - ECG_m)}$
	$HRV_{SD} = \sqrt{\frac{1}{N}\sum(RR(n) - HRV_m)}$
Shannon entropy	$ECG_e = \sum P(ecg(n)) \cdot \log(P(ecg(n)))$,
	$HRV_e = \sum P(RR(n)) \cdot \log(P(RR(n)))$

By following the procedure for managing the time-windows and organizing the training dataset, the vector of features contains the mean (m_k), standard deviation (*std*), and Shannon entropy (sh_e_k) for one to five minutes, a total of 15 features for each observation as shown in Fig. 4.

$$\begin{aligned} f_n = [&m_1,\ std_1,\ sh_e_1,\ m_2,\ std_2,\ sh_e_2,\ m_3,\\ &std_3,\ sh_e_3,\ m_4,\ std_4,\ sh_e_4,\ m_5,\ std_5,\ sh_e_5] \end{aligned} \tag{1}$$

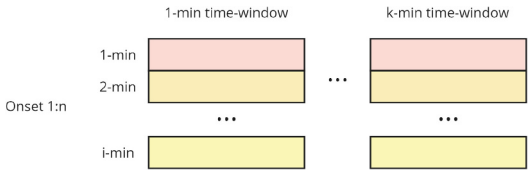


Fig. 4. Training dataset.

2.3 Algorithm Experimentation

A Matlab application developed for this research produces the training and testing datasets which are inputs to classification algorithms, producing accuracy, precision, recall, and F1 score metrics as shown in Table 2, which are calculated from the confusion matrix in Fig. 5. TP refers to the number of positive observations classified correctly, TN to the number of negative observations classified correctly, while FP and FN number of positive and negative observations classified incorrectly, it is, negative and positive, respectively.

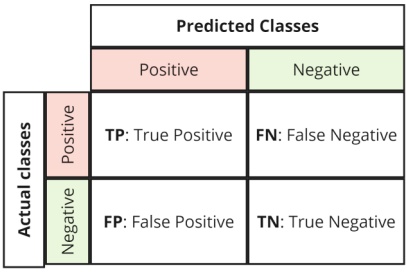


Fig. 5. Confusion matrix.

In this phase, different experiments are made to test the performances of the classification algorithms and validate the initial hypothesis. Some of the experiments evaluated in this research include:

- Applying different classification techniques, such as support vector machine, random forest, and k-nearest neighbors.
- Testing the same classification technique with an extended set of features, that correspond to time-domain features on longer time-windows.
- Comparing performance between algorithms trained with the Shannon entropy as feature, and algorithms trained with mean and standard deviation and time domain features.
- Comparing performance between using the ECG or the HRV signals as a source for feature extraction.

2.4 Evaluation

The last phase consists of comparing the performance metrics in Table 3, each experiment and analyzing the scenarios where the prediction reaches the highest accuracy. This phase involves reviewing all metrics to determine the extent to which a lower, yet still acceptable, accuracy or a specificity farther from the onset is feasible. Additionally, this phase includes a discussion of the reasons why the proposed method achieves competitive metrics.

Table 3. Algorithm performance metrics.

Recall	*Precision*
$\frac{TP}{TP+FN}$	$\frac{TP}{TP+FP}$
Accuracy	*F1Score*
$\frac{TP+TN}{TP+TN+FP+FN}$	$\frac{2 \cdot precision \cdot recall}{precision+recall}$

- A testing dataset was created for each minute from 1 to 15 min before the onset, to evaluate the algorithm performance over a longer period as shown in Fig. 6.

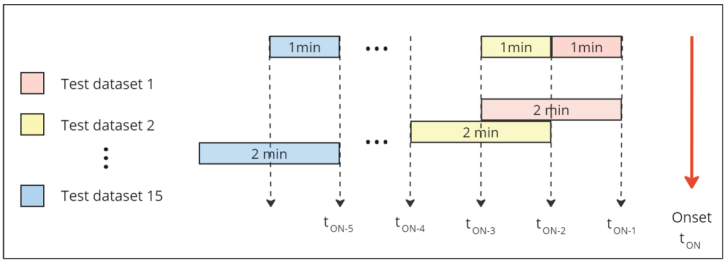


Fig. 6. Testing dataset creation.

3 Results

The Table 4 shows the results for three classification algorithms trained with the Shannon entropy as the only feature, it is, $f_n = [sh_e_1, sh_e_2, sh_e_3, sh_e_4, sh_e_5]$. The convolutional neural network, k-nearest neighbors, decision trees, and support vector machine algorithms are already reported in the literature and were taken as references for this study as they were the methods with highest performance in previous studies. SVMs are suitable for classifying time series data [32] and use efficiently the memory in a real-time scenario, while CNNs are able to learn features directly from the raw data detecting patterns in different time scales.

Table 4. Performance metrics.

Classification	# minutes before the onset	Accuracy (%)	F1 score (%)
Convolutional neural network	1	92.5	92.31
	5	97.5	97.44
K-nearest neighbors	1	90	89.47
	5	90	90.48

(*continued*)

Table 4. (*continued*)

Classification	# minutes before the onset	Accuracy (%)	F1 score (%)
Binary decision trees	5	97.5	97.44
	1	85	83.33
Support vector machine	1	90	89.47
	5	97.5	97.44

Table 1 shows relevant results from the literature, but these papers are limited to present the prediction metrics for one or five minutes before the onset. Results for this method in Table 4 follow this rule to ease the comparison between results, but detailed plots are also shown for the minute-to-minute testing dataset.

The first experiment run with the training and datasets created aimed to evaluate the impact of using multiple length time-windows. Thus, a first comparison was made between two sets of features: $f_n = [sh_e_1]$ and $f_n = [sh_e_1, sh_e_2, sh_e_3, sh_e_4, sh_e_5]$. Results for the accuracy and F1 Score are shown in Fig. 7, where each point in the plot corresponds to each point in the testing dataset presented in Fig. 6, it is, metrics for each minute from 1 to 15 min before the cardiac event.

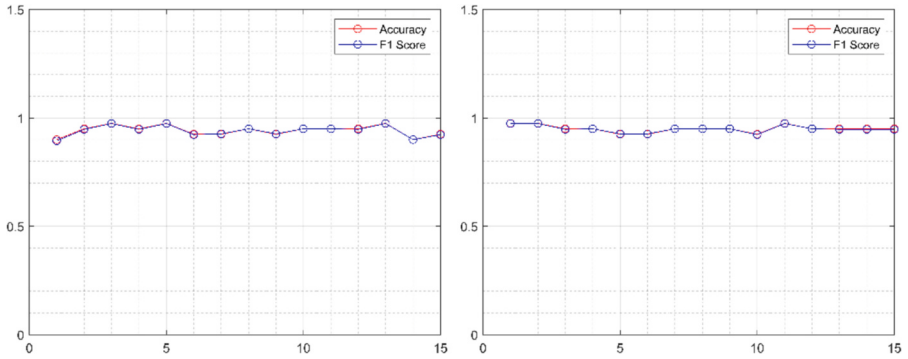


Fig. 7. Metrics for classification using the Shannon entropy as a feature.

A similar experiment is made using traditional time-domain features, the mean, and the standard deviation extracted from ECG signals. Figure 8 shows the results for $[m_1, std_1]$ at the left, and $f_n = [m_1, std_1, m_2, std_2, m_3, std_3, m_4, std_4, m_5, std_5]$ at the right.

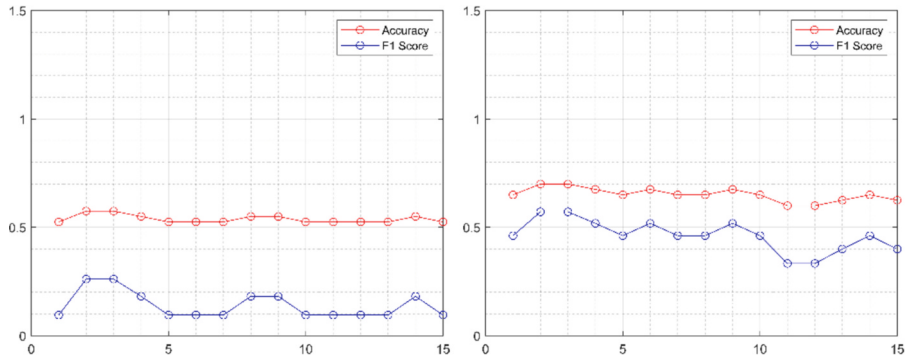


Fig. 8. Metrics for classification using mean and standard deviation as features.

Another experiment was run using the HRV signal as source to create the dataset. Results are shown in Fig. 9.

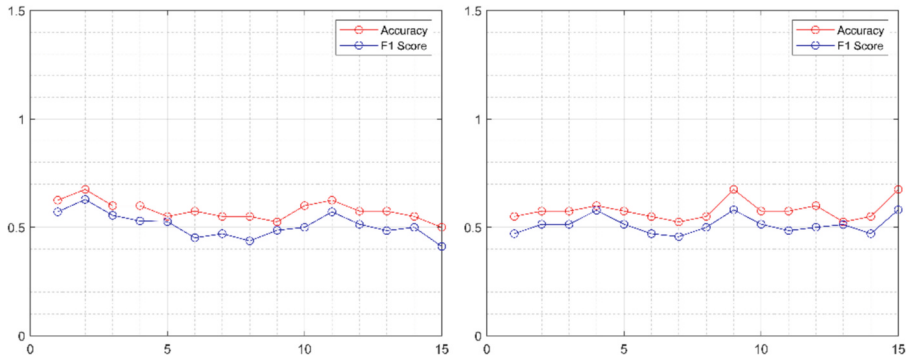


Fig. 9. Metrics for Shannon entropy extracted from HRV signal.

4 Discussion

The results from the experiments show three significant facts regarding the approach used in this research for each phase of the method, the use of Shannon entropy as a feature, organizing time-windows for extracting the features, and testing for each minute over a longer time before the onset.

Classification with the Shannon entropy as the unique feature provides metrics near 90%, which is considered as a good performance. Results for mean and standard deviation are consistent with the literature and work as a reference for the experiments. As stated earlier, results for this method (Table 4) allow the comparison in the same terms as the literature, from this, CNN and SVM algorithms get the highest values and are comparable with results that use the same databases for training the prediction algorithm. Also, for comparison purposes, Table 5 shows the results from this and previous works, pointing out the diversity of times and methods that have been tested.

Table 5. Results comparison.

Year	Title		Results
2018	Ebrahimzadeh [14] 23 features	13 min	Accuracy: 90.18%
2020	Tseng [33] 2D frequency domain features CNN classifier	5 min	Accuracy: 88%
2021	Parsi [19] 6 features k-NN classifier	5 min	Accuracy: 91.5%
2022	Manhong [34] 5 features SVM classifier	4 min	Accuracy: 97.9%
2023	Karimulla [35] 77 features	60 min	Accuracy: 96.43%
2024	Proposed method 1 feature	15 min	Accuracy: 92.5%

While several of the results are obtained extracting features from multiple domains, and training the model with more than 20 features, this research uses only the Shannon entropy and gets high values of both accuracy and F1 Score. Implications in a real-time scenario are that the automatic detection will require less memory and arithmetic operations, making it feasible to implement the model on wearable devices, this means that hardware will require fewer resources for predicting the cardiac event, and the patient has a lower risk of getting false positives detections, meaning an undesired electric shock, in the case of ICD users.

An improvement of the metrics when using information from more minutes before the onset, both plots increase the values for accuracy and F1 score on the extended feature set. The increment becomes higher for mean and standard deviation (Fig. 6), the accuracy goes from 52% to 63% approximately, while for the Shannon entropy it goes from 90% to 95% (Fig. 5), taking the result for one minute before the onset. Arranging the datasets in a way that extends the set of features from a unique feature can be applied to existing prediction techniques that use a large number of features from multiple domains, allowing the identification of the most relevant. In the results obtained in this research, the improvements were more evident for a technique that traditionally has poor performance. While using mean and standard deviation results in an accuracy comparable to a random prediction, and high probabilities of both false positives and negatives, extending the feature set with data from longer time-windows, increases accuracy, but mostly, it decreases the chances of getting false positives or failing in detect a real onset.

As expected, the performance decreases as the predictions are made earlier from the onset, evidencing that the amount of information about the occurrence of the cardiac

event increases as the event becomes imminent. However, with this method the prediction is still possible as far as 15 min before the onset, and metrics don't show a rapid decay.

Limitations in this study are related mainly to the amount of data available for constructing large datasets. Early prediction needs to be made over a long enough time-series of ECG data, and Physionet has few databases with long-term records. For getting a larger dataset, multiple cardiac events are identified on the same record, and data is extracted if is separated at least 30 min from a prior event. This is also a limitation, because the ECG records come from patients in ICU and keep an altered rhythm over several hours, making the segments alike. Applying the results of this work to a permanent monitor as a wearable device, probably won't achieve the same performance, and it will require the gathering of more data to build a more customized model. As Physionet databases include records from specific populations, work with local databases will be necessary to fit the results to patients with different characteristics.

5 Conclusions and Future Work

The current findings of the research can be summarized in the following points (1) Adding longer time-windows as additional features, improves the predictor's performance. (2) Shannon entropy performs well as feature on ECG signals, indicating that information about the Cardiac event exists at least 12 min before the onset. (3) HRV signal does not carry as much information as ECG does. For real-time applications it would not be enough to know the heart rate. (4) The model's performance is coherent with the literature's results, as SVM keeps the highest metrics in prediction. The experiments presented in this ongoing research show promising results for the method proposed to predict an SCD onset. A 0.95 F1 score ten minutes before the onset using the Shannon entropy as unique feature reaches performances already reported that use significantly more features from time and frequency domain, simplifying the prediction process and easing its hardware implementation. This becomes evidence of the existence of enough information in an early ECG signal, and its nature is better interpreted through the Shannon entropy. The question about why the Shannon entropy is a better descriptor points to the fact that the heartbeat, as measured electrically and graphed respect time, has a standard form or morphology that can be considered "ordered". When this morphology is altered because of the near happening of a heart failure, the following samples will become more difficult to predict, changing drastically the value of the entropy, which is a measure of the sample's predictability. This is reinforced by the fact that training the algorithm with the Shannon entropy extracted for the HRV signal, which has just a sample per heartbeat, results in a poor performance, leading to the conclusion that a full set of samples carries more information related to the onset occurrence (Fig. 7). Future work of this research follows two directions, the first one is related to develop another series of experiments in which the internal parameters are fine-tuned. This work used basic and default model's architecture (e.g. gaussian SVM with 5 cross-fold validation) to compare performance between different algorithms but working with variations of the same classifier can lead to additional improvements in performance. Another line of work deals with implementation issues in wearable devices, therefore works can be proposed to evaluate embedded microprocessors running all the processes needed to make a real-time prediction, including heart rate calculation, filtering, time windows management,

Shannon entropy extraction, and onset prediction. A first approach will use an existing ECG monitor with wearable specifications, that allows permanent measurements over long periods of time for building a different and new dataset to apply the prediction algorithm.

References

1. Krokhaleva Y, Vaseghi M (2018) Update on prevention and treatment of sudden cardiac arrest. https://doi.org/10.1016/j.tcm.2018.11.002
2. Fox CS, Evans JC, Larson MG, Kannel WB, Levy D (2004) Temporal trends in coronary heart disease mortality and sudden cardiac death from 1950 to 1999: The Framingham Heart Study. Circulation. https://doi.org/10.1161/01.CIR.0000136993.34344.41
3. Ebrahimzadeh E, Pooyan M, Bijar A (2014) A novel approach to predict sudden cardiac death (SCD) using nonlinear and time-frequency analyses from HRV signals. PLoS One 9. https://doi.org/10.1371/JOURNAL.PONE.0081896
4. Wong CX et al (2019) Epidemiology of Sudden cardiac death: global and regional. Perspectives. https://doi.org/10.1016/j.hlc.2018.08.026
5. Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M, Dai S, De Simone G, Ferguson TB, Ford E, Furie K, Gillespie C, Go A, Greenlund K, Haase N, Hailpern S, Ho PM, Howard V, Kissela B, Kittner S, Lackland D, Lisabeth L, Marelli A, McDermott MM, Meigs J, Mozaffarian D, Mussolino M, Nichol G, Roger VL, Rosamond W, Sacco R, Sorlie P, Stafford R, Thom T, Wasserthiel-Smoller S, Wong ND, Wylie-Rosett J (2010) Executive summary: Heart disease and stroke statistics-2010 update: a report from the American heart association. https://doi.org/10.1161/CIRCULATIONAHA.109.192667
6. Byrne R et al (2008) Multiple source surveillance incidence and aetiology of out-of-hospital sudden cardiac death in a rural population in the West of Ireland. Eur Heart J. https://doi.org/10.1093/eurheartj/ehn155
7. Chugh SS, Jui J, Gunson K, Stecker EC, John BT, Thompson B, Ilias N, Vickers C, Dogra V, Daya M, Kron J, Zheng ZJ, Mensah G, McAnulty J (2004) Current burden of sudden cardiac death: Multiple source surveillance versus retrospective death certificate-based review in a large U.S. community. J Am Coll Cardiol. https://doi.org/10.1016/j.jacc.2004.06.029
8. De Vreede-Swagemakers JJM et al (1997) Out-of-hospital cardiac arrest in the 1990s: a population-based study in the Maastricht area on incidence, characteristics and survival. J Am Coll Cardiol. https://doi.org/10.1016/S0735-1097(97)00355-0
9. Jones JL, Tovar OH (1996) The mechanism of defibrillation and cardioversion. Proc IEEE. https://doi.org/10.1109/5.486742
10. Kaplan Berkaya S, Uysal AK, Sora Gunal E, Ergin S, Gunal S, Gulmezoglu MB (2018) A survey on ECG analysis. Biomed Signal Process Control 43:216–235. https://doi.org/10.1016/j.bspc.2018.03.003
11. Boveda S, Garcia R, Defaye P, Piot O, Narayanan K, Barra S, Gras D, Providencia R, Algalarrondo V, Beganton F, Perier MC, Jacob S, Bordachar P, Babuty D, Klug D, Leclercq C, Fauchier L, Sadoul N, Deharo JC, Marijon E (2019) Implantable cardioverter defibrillator therapy for primary prevention of sudden cardiac death in the real world: main findings from the French multicentre DAI-PP programme (pilot phase). https://doi.org/10.1016/j.acvd.2019.05.005
12. Faust O, Hong W, Loh HW, Xu S, Tan R-S, Chakraborty S, Barua PD, Molinari F, Acharya UR (2022) Heart rate variability for medical decision support systems: a review. Comput Biol Med 145:105407. https://doi.org/10.1016/j.compbiomed.2022.105407

13. Parsi A, O'Loughlin D, Glavin M, Jones E (2020). Prediction of sudden cardiac death in implantable cardioverter defibrillators: a review and comparative study of heart rate variability features. https://doi.org/10.1109/RBME.2019.2912313
14. Ebrahimi Z, Loni M, Daneshtalab M, Gharehbaghi A (2020) A review on deep learning methods for ECG arrhythmia classification. Expert Syst with Appl. X. 7:100033. https://doi.org/10.1016/j.eswax.2020.100033
15. Fujita H et al (2016) Sudden cardiac death (SCD) prediction based on nonlinear heart rate variability features and SCD index. Appl Soft Comput J 43:510–519. https://doi.org/10.1016/j.asoc.2016.02.049
16. Amezquita-Sanchez JP, Valtierra-Rodriguez M, Adeli H, Perez-Ramirez CA (2018) A novel wavelet transform-homogeneity model for sudden cardiac death prediction using ECG signals. J Med Syst. https://doi.org/10.1007/s10916-018-1031-5
17. Ebrahimzadeh E, Manuchehri MS, Amoozegar S, Araabi BN, Soltanian-Zadeh H (2018) A time local subset feature selection for prediction of sudden cardiac death from ECG signal. Med Biol Eng Comput. https://doi.org/10.1007/s11517-017-1764-1
18. Ebrahimzadeh E et al (2019) An optimal strategy for prediction of sudden cardiac death through a pioneering feature-selection approach from HRV signal. Comput Methods Programs Biomed. https://doi.org/10.1016/j.cmpb.2018.12.001
19. Parsi A, Byrne D, Glavin M, Jones E (2021) Heart rate variability feature selection method for automated prediction of sudden cardiac death. Biomed Signal Process Control 65:102310. https://doi.org/10.1016/j.bspc.2020.102310
20. Mandal S, Mondal P, Roy AH (2021) Detection of Ventricular Arrhythmia by using Heart rate variability signal and ECG beat image. Biomed. Signal Process Control 68:102692. https://doi.org/10.1016/j.bspc.2021.102692
21. Yang J, Sun Z, Zhu W, Xiong P, Du H, Liu X (2023) Intelligent prediction of sudden cardiac death based on multi-domain feature fusion of heart rate variability signals. EURASIP J Adv Signal Process. https://doi.org/10.1186/S13634-023-00992-6
22. Lai D, Zhang Y, Zhang X, Su Y, Bin Heyat MB (2019) An automated strategy for early risk identification of sudden cardiac death by using machine learning approach on measurable arrhythmic risk markers. IEEE Access 7:94701–94716. https://doi.org/10.1109/ACCESS.2019.2925847
23. Marinho LB, Nascimento N de MM, Souza JWM, Gurgel MV, Rebouças Filho PP, de Albuquerque VHC (2019) A novel electrocardiogram feature extraction approach for cardiac arrhythmia classification. Futur Gener Comput Syst 97:564–577. https://doi.org/10.1016/j.future.2019.03.025
24. Ebrahimzadeh E, Pooyan M (2011) Early detection of sudden cardiac death by using classical linear techniques and time-frequency methods on electrocardiogram signals. J Biomed Sci Eng. https://doi.org/10.4236/jbise.2011.411087
25. Ebrahimzadeh E, Pooyan M, Bijar A (2014) A novel approach to predict sudden cardiac death (SCD) using nonlinear and time-frequency analyses from HRV signals. PLoS ONE. https://doi.org/10.1371/journal.pone.0081896
26. Panjaitan F, Nurmaini S, Akbar M, Mirza AH, Syaputra H, Kurniawan TB, Umi RP (2019) Identification of Classification Method for Sudden Cardiac Death: A Review. In: ICECOS 2019—3rd international conference on electrical engineering and computer science, Proceeding. Institute of Electrical and Electronics Engineers Inc. pp 93–97. https://doi.org/10.1109/ICECOS47637.2019.8984465
27. Goldberger AL et al (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 101:215–220. https://doi.org/10.1161/01.cir.101.23.e215
28. Greenwald SD (1986) The development and analysis of a ventricular fibrillation detector

29. Sudden Cardiac Death Holter Database v1.0.0
30. MIT-BIH Normal Sinus Rhythm Database v1.0.0
31. Silva I, Moody GB (2014) An open-source toolbox for analysing and processing physionet databases in MATLAB and Octave. J Open Res Softw. 2. https://doi.org/10.5334/JORS.BI
32. Burges CJC (1998) A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov 2:121–167. https://doi.org/10.1023/A:1009715923555/METRICS
33. Tseng LM, Tseng VS (2020) Predicting ventricular fibrillation through deep learning. IEEE Access. 8:221886–221896. https://doi.org/10.1109/ACCESS.2020.3042782
34. Shi M, Yu H, Wang H (2022) Automated detection of sudden cardiac death by discrete wavelet transform of electrocardiogram signal. Symmetry 14:571. https://doi.org/10.3390/SYM14030571
35. Karimulla S, Patra D (2024) An optimal methodology for early prediction of sudden cardiac death using advanced heart rate variability features of ECG signal. Arab J Sci Eng 49:6725–6741. https://doi.org/10.1007/S13369-023-08457-6/METRICS

Facial Weakness Detection in Stroke Using Feedback Cascade Regression and Support Vector Machine

Mohamad Shavin(✉), Tri Arief Sardjono, and Nada Fitrieyatul Hikmah

Department of Biomedical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
mohamadshavin19@gmail.com

Abstract. Stroke is a leading cause of death and disability worldwide. Early detection of stroke symptoms, such as facial weakness, is critical for timely treatment. The FAST method (Facial drooping, Arm weakness, Speech difficulty, and Time) is widely used for identifying these symptoms. This study introduces a novel system for detecting facial weakness in stroke patients using Feedback Cascade Regression for facial landmark alignment and Support Vector Machine (SVM) for classification. The proposed method calculates 29 facial tilt features to assess facial symmetry and employs Principal Component Analysis (PCA) for dimensionality reduction. The SVM model, optimized with a grid search for parameters (C = 100, gamma = 0.1), achieved a 95% accuracy in distinguishing between stroke-affected and normal faces. The findings suggest that integrating advanced pre-processing techniques with SVM classification can significantly enhance facial weakness detection. This research contributes to improving the accuracy and robustness of stroke detection systems. Implications: The proposed system can be integrated into clinical practice for early stroke detection, potentially improving patient outcomes through timely intervention.

Keywords: Facial weakness · Feedback cascade regression · Stroke · SVM

1 Introduction

Stroke is the second leading cause of death and the third leading cause of disability. One in four people is at risk of experiencing a stroke in their lifetime. About 87% of stroke-related deaths and disability-adjusted life years are significant concerns [1]. Despite this, many stroke cases are not quickly recognized, making awareness of its symptoms crucial for prevention and further treatment. Increasing public literacy using the FAST method (Facial drooping, arm weakness, and speech difficulty) is a practical guideline for identifying the main symptoms of a stroke. One symptom related to the face is the inability to move part of it. This is often visible when a person smiles or is at rest, with signs such as drooping eyes, asymmetrical eyebrows, or a drooping face [2, 3].

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 235–249, 2026.
https://doi.org/10.1007/978-981-95-9233-3_18

Stroke-induced facial weakness results in pronounced asymmetry, often affecting one side of the face more than the other, and can impair movements such as smiling, raising eyebrows, or closing eyes. Current clinical practice uses screening tools like pen-and-paper checklists, but their accuracy and reliability are limited due to variability in training and experience of non-neurologist providers. Identifying neurological diseases such as facial weakness requires comprehensive expertise [4, 5]. Furthermore, stroke symptoms may manifest as sudden numbness or weakness in the face, arm, or leg, particularly on one side of the body. Additional indicators can encompass confusion, challenges in understanding speech, sudden vision disturbances in one or both eyes, difficulty walking, dizziness, loss of balance or coordination, and abrupt, severe headaches with no discernible cause. Prompt recognition of these symptoms and seeking immediate medical attention are vital to reducing the long-term impact of a stroke.

Numerous studies have employed machine learning techniques for medical image processing in disease detection, including the use of Support Vector Machine (SVM) [6] with feature extraction, as well as convolutional neural networks [7, 8]. The study by Zhuang et al. [9] explored the use of digital images for the timely detection of facial weakness. It evaluated various methods for measuring facial weakness, including landmark-based and intensity features, and found that approaches using the Histogram of Oriented Gradients (HoG) feature were generally more accurate in detecting facial weakness from a single static image. This study employed the Support Vector Machine (SVM) method, utilizing a linear SVM classifier that combines facial landmark features with HoG features to enhance classification performance in detecting facial weakness. Other studies have adopted different methodologies. For instance, Hao et al. [10] proposed a feedback cascade of regression method, which involves a stepwise regression process from significant facial landmarks to all landmarks, along with feedback operations to improve regression robustness. This approach began by identifying five significant facial landmarks, such as the centers of the eyes, the tip of the nose, and the corners of the mouth, using a rough shape regression model. Additionally, Dominguez et al. [11] proposed using key point analysis to detect facial paralysis in photos. This method involved three modules: facial landmark extraction, facial size calculation, and facial paralysis classification. The facial size calculation process measured the distances between facial landmarks to determine the spatial relationships between them. The proposed facial size aimed to identify the level of asymmetry between the left and right sides of the face and describe facial gestures for input into a multilayer perceptron. Despite the advancements in these studies, there are limitations in the pre-processing stage, particularly concerning the placement of facial points. The commonly used cascade method often encounters errors in the placement of facial points under various conditions or expressions, which can impact detection accuracy.

The research conducted by Zhuang et al. used a face detection method with a standard cascade approach, which, although effective in certain conditions, remains highly sensitive to variations in facial conditions. This sensitivity can lead to errors in detecting facial landmarks, which impacts the accuracy of facial weakness detection. To address this limitation, this study developed an approach using Feedback Cascade Regression. According to Hao et al., the feedback cascade method is superior in detecting faces and is more adaptive to various facial conditions, thereby reducing errors in landmark placement and improving detection accuracy.

In this study, a stroke symptom recognition system is proposed that utilizes a two-class classification approach to distinguish between stroke patients with facial weakness and healthy individuals. Support Vector Machine (SVM) is employed for classification, constructing a model that separates data into different classes using the optimal hyperplane. The study aims to enhance the pre-processing, feature extraction, and in-depth analysis stages to achieve definitive results. To address the issue of facial alignment, the Feedback Cascade of Regressions (FCR) method is proposed. The FCR method is superior to the common cascade method as it helps align facial images by accurately detecting the positions of facial landmarks. During the feature extraction stage, the study uses 29 facial tilt calculations to assess the degree of symmetry between the two sides of the face. Subsequently, the extracted features undergo dimensionality reduction using Principal Component Analysis (PCA) [12] to minimize the number of features in the dataset while retaining as much information as possible. The PCA-reduced features are then classified using the SVM method. SVM operates by identifying the optimal hyperplane that separates data into distinct classes with the largest margin. It is effective in handling high-dimensional data and can be used with various kernels to address non-linear problems [13]. To evaluate the model's performance, the K-Fold Cross Validation method is employed [14]. The combination of PCA and SVM aids in reducing the dimensionality of the data while preserving relevant information for the classification task.

This study proposes a stroke symptom recognition system that utilizes Feedback Cascade Regression (FCR) for facial alignment and Support Vector Machine (SVM) for classification. By addressing the limitations in pre-processing and landmark placement, this research aims to improve the accuracy and robustness of facial weakness detection in stroke patients. The contribution of this study lies in its novel approach to combining advanced pre-processing techniques with machine learning models, offering a more reliable tool for early stroke detection.

2 Materials and Method

This study builds on several theoretical foundations that are critical to the proposed stroke symptom recognition system. The key methods include Data preparation, Pre-processing and Normalization, Face Detection, Landmarking, and Alignment, Feature Extraction, Principal Component Analysis (PCA), and Support Vector Machine (SVM).

2.1 Data Preparation

The Toronto NeuroFace Dataset is specifically designed to support research in detecting and classifying facial weakness, particularly within the context of neurological disorders such as stroke and Amyotrophic Lateral Sclerosis (ALS). This dataset was compiled by Dr. Yana Yunusova and her team at the UHN-Toronto Rehabilitation Institute and Sunnybrook Research Institute, with financial support from the Michael J. Fox Foundation, NIH-NIDCD, and the Natural Sciences and Engineering Research Council of Canada. It comprises 261 annotated videos of individuals with various neurological conditions and healthy control subjects (HC). Over 3300 video frames have been manually annotated by clinical experts using 68 facial landmarks, which include key features such as the eyes, nose, mouth, and facial edges. These annotations are instrumental in assessing facial symmetry and identifying facial weakness [15]. This study utilizes the Toronto NeuroFace Dataset, which encompasses videos of individuals with various neurological conditions and healthy control subjects (HC). Out of the 3300 available image frames, 476 images were selected for analysis, comprising 236 images from HC (normal) subjects and 237 images from stroke patients exhibiting a range of conditions.

2.2 General Systems Design

The system design to be implemented comprises several stages. The code will be developed using Visual Studio Code (VSC) to address the research objectives. Initially, the Toronto NeuroFace dataset will be utilized. After collecting the dataset, facial images will undergo pre-processing, which includes cropping, resizing, rescaling, grayscale conversion, and rotation. The processed images will then be subjected to landmarking, wherein facial landmarks are identified and recorded at specific points on the face. Key features such as the eyes, eyebrows, and mouth will be extracted. Subsequently, Regions of Interest (ROI) will be selected, focusing on areas such as the eyes, nose, eyebrows, mouth, and the entire face. Following landmark detection, face alignment will be performed using the Feedback Cascade of Regressions (FCR) method to refine the placement of facial landmarks. Once the landmarks are accurately positioned, the feature extraction process will calculate 29 facial symmetry metrics. The extracted data will be analyzed using Principal Component Analysis (PCA) to reduce dimensionality and extract relevant parameters for the classification process. The final stage involves classification using Support Vector Machine (SVM) with four different kernels and various parameters, resulting in two categories: normal and stroke. Post-classification, the results will be compared with other classifiers, and K-Fold Cross Validation will be employed to assess the model's generalization capability. Finally, performance evaluation and confusion matrix analysis will be conducted. The block diagram for this research can be seen in Fig. 1.

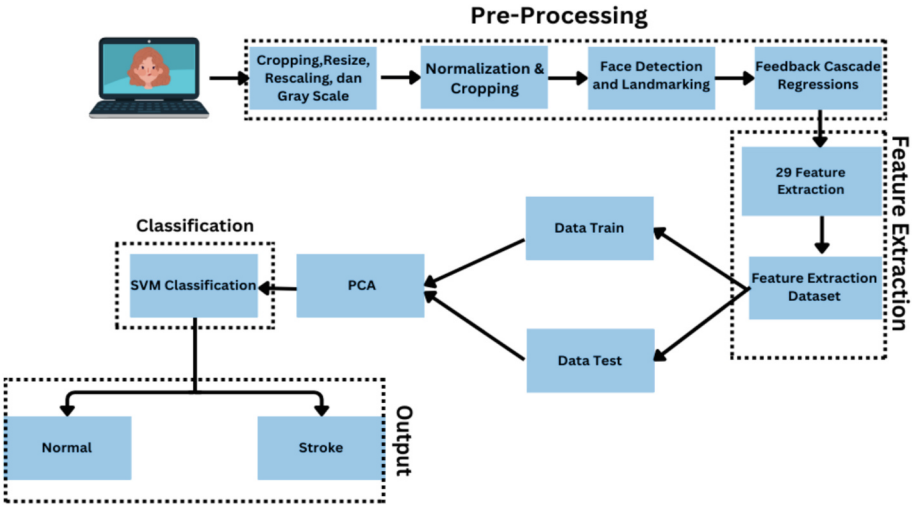


Fig. 1. Block Diagram Systems

2.3 Pre-processing and Normalization

The pre-processing methods standardize the collected data to ensure uniform processing and consistent results [16]. Techniques include resizing, rescaling, grayscale conversion, and cropping, with resizing typically adjusting image dimensions to 128x128 pixels. Normalization involves rotating the image based on the angle calculated from the difference between the top right and left jaw points, ensuring the face is properly aligned for further analysis. The normalization process is carried out using image rotation, where the rotation angle is determined by calculating the difference in position between the top right and top left jaw points. The degree of inclination can be calculated using Eqs. (1), (2), and (3) [17, 18].

$$\angle(\text{pa}, Pb) = \arctan 2\ (\Delta x,\ \Delta y) x 180/\pi \tag{1}$$

$$y = Pay - Pby \tag{2}$$

$$\Delta x = Pax - Pbx \tag{3}$$

where Δx is the difference in the x-axis between the two points on the left and right upper jaw, and Δy is the difference in the y-axis between the two points on the left and right upper jaw. Pa and Pb are feature points on the face.

2.4 Face Detection, Landmarking, and Alignment

Face detection uses HOG (Histogram of Oriented Gradients) and SVM (Support Vector Machine) in Dlib. The image is converted to grayscale, and HOG calculates gradient histograms to detect faces using a sliding window approach, with an SVM classifying

whether a window contains a face. Once a face is detected, landmark detection begins using Dlib's regression model, which refines facial landmark estimates through cascade regression trees until 68 landmarks are accurately placed [19–21]. Face alignment involves refining these landmarks using feedback cascade regression, which improves accuracy through three stages: salient, inner, and all. This method ensures precise placement of facial landmarks, which is crucial for feature extraction [10]. Feedback Cascade Regression method progressively calculates the location of landmarks on the face. Equation (4) illustrates the process:

$$S_{i,t+1} = S_{i,t} + r_t\left(I, S_{i,t}\right) \tag{4}$$

where *rt* represents the t-th regressor, *S*i,*t* + 1 represents the shape at the next level, and the shape is updated step by step, with the difference at the next level denoted by *rt*(*I*, *S*i,*t*). Following this, a feedback strategy is implemented after detecting all landmarks. The positions of key landmarks are updated with the detected results of all landmarks. By replacing rough initial landmark positions with more accurate results, this process provides a new starting point approaching global optimum, enhancing the accuracy of key landmarks, and ultimately improving the final result. Equation (5) demonstrates this adjustment process:

$$x_{i+1} = \begin{bmatrix} X_i + 11 \\ X_i + 12 \\ X_i + 1_p \end{bmatrix} = \begin{bmatrix} X_i + 11 \\ X_i + 12 \\ X_i + 1_p \end{bmatrix} = \begin{bmatrix} Dx_{i1} \\ Dx_{i2} \\ Dx_{ip} \end{bmatrix} \tag{5}$$

where x_{i+1} ix the pixel vector at the i + 1-th level, containing pixel value at index i in the pixel vector, D denotes the number of pixel in vector i + 1. The landmark extraction method is a necessary step in facial recognition to determine the presence of important facial features required for face detection. To extract facial landmarks from a face image, an algorithm is needed to detect the region of interest (ROI) on the face, such as the eyes and mouth. To obtain the desired landmarks, we can extract the detected points within the desired ROI [22, 23].

2.5 Feature Extraction

In the feature extraction stage, parameters are derived for the subsequent classification process using SVM. After normalization and face alignment with feedback cascade regressions, facial landmarks are accurately positioned, enabling effective feature extraction. This stage measures various facial points, assessing aspects such as curvature and slope. This stage measures various facial points, such as the curvature or slope between points, using Eqs. (7), (8), and (9), and parameters from nearby segments are measured using Equation:

$$m(Pa, Pb) = \left| \frac{Pay - Pby}{Pax - Pbx} \right| \tag{6}$$

$$d(Pa, Pb) = \sqrt{(Pax - Pbx)^2 + (Pay - Pby)}^2 \tag{7}$$

$$\overline{S}(Ps, \ldots, P1) = \sum_{x-s}^{I-1} d(Px, Px+1) + d(Ps, P1) \tag{8}$$

where $\overline{S}$, represents a nearby segment Ps is the starting point of the segment, and Pl is the endpoint of the segment. The assumption is that a normal face exhibits symmetry between its left and right sides; asymmetry in these facial parts may indicate a stroke [11]. Utilizing 29 symmetry measurement values on the face, the MEEshape predictor, a publicly available shape predictor, identifies 68 points, with 51 points being critical for detecting facial paralysis [24–26].

2.6 Performance Evaluation of PCA

Principal Component Analysis (PCA) is a dimensionality reduction method that transforms correlated features into uncorrelated ones, effectively reducing the complexity of the data while preserving as much variance as possible. By projecting data into a lower-dimensional space, PCA helps simplify the dataset, making it more manageable for analysis and improving the performance of subsequent machine learning models. This technique is particularly useful for handling high-dimensional data and mitigating the risk of overfitting [12, 27].

2.7 Support Vector Machine

Support Vector Machines (SVM) classify data by finding the optimal hyperplane that separates different classes with the maximum margin. The algorithm identifies support vectors, the critical data points closest to the separating hyperplane, which influence its position [28]. SVM can handle non-linear data by utilizing various kernel functions, such as Gaussian/Radial Basis Function (RBF), sigmoid, linear, and polynomial kernels. These kernels transform the input space into a higher-dimensional space, allowing SVM to construct a linear boundary in this transformed space. Key parameters for SVM include the regularization parameter (C) and gamma (γ), which help balance the trade-off between achieving a low training error and a low testing error [29, 30]. Kernel RBF (Radial Basis Function) is commonly used when dealing with data that is difficult to separate linearly, as shown in Eq. (9):

$$K(x_i, y_i) = \exp\left(-\frac{xi - yi^2}{2\sigma^2}\right) \tag{9}$$

3 Result

In this chapter, the results of the conducted tests are presented. This section discusses the processes of data collection and preparation, image processing, and the classification of processed images into their respective classes: normal and stroke. The program model is developed using Visual Studio Code with Python.

3.1 Pre-Processing

In this pre-processing stage, various steps are necessary due to the highly varied dataset collected from different sources. These steps facilitate the subsequent methods and are illustrated using original face images of the author, as shown in Fig. 1. Initially, the face images are resized to a target size of 128x128 pixels without altering the aspect ratio; the width or height is adjusted accordingly to maintain the aspect ratio. The rescale process then centers the resized face within the 128x128 pixels frame by adding proportional borders. Next, normalization adjusts the pixel values to fall within the range [0, 1] to improve numerical stability and model performance in machine learning algorithms. The resized and bordered face image's data type changes to floating-point format, and each pixel value is divided by 255. Image rotation adjusts the face image's orientation based on the tilt angle, calculated from the tilt difference between two jawline endpoints. Finally, the grayscale method converts colored face images into grayscale images by removing color information and retaining only luminance intensity, simplifying the images for further analysis.

3.2 Detection and Alignment of Facial Landmarks Using Feedback Cascade

Face detection begins with the dlib library to identify the face and determine 68 facial landmarks, including features such as the contour, eyebrows, eyes, nose, and mouth.

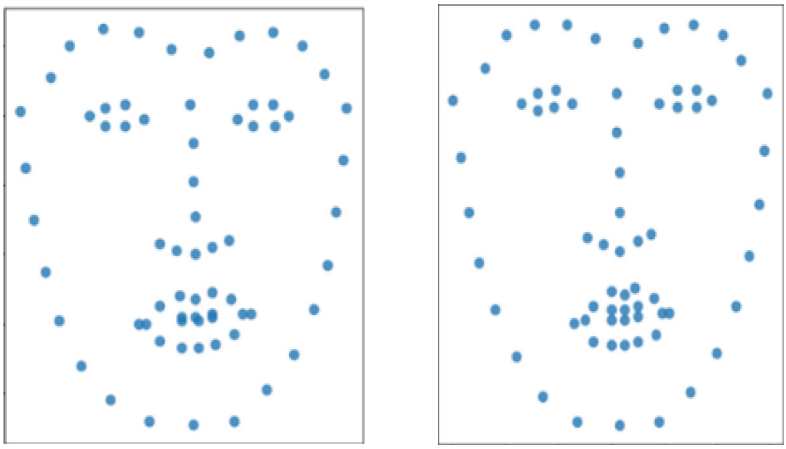


Fig. 2. The Results of the Analysis Process: Part (a) shows the face detection and landmarking process using the dlib library. Part (b) shows the results after the feedback cascade process. This figure demonstrates the implementation of the facial landmark alignment technique used in the research.

Face alignment is achieved using feedback cascade regression, which involves an iterative, three-stage approach: detecting key landmarks (salient stage), more detailed features (inner stage), and all facial landmarks (all stage). Cascade regression updates the

landmark positions iteratively, refining the shapes based on predictions from the regressor. A feedback strategy further adjusts the landmark positions to improve accuracy. The results of this alignment process are illustrated in Fig. 2.

3.3 Feature Extraction

After the facial landmarks are aligned, 29 facial points are measured to determine the symmetry of the facial regions. The measurement of these 29 facial points. By comparing the symmetry levels between the left and right sides of the face, a value between 0 and 1 is obtained, where a value of 0 is assumed to be asymmetrical (Stroke) and a value of 1 is assumed to be nearly symmetrical (Normal).

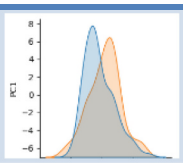
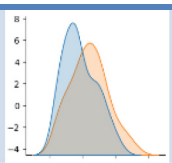
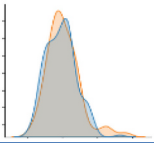
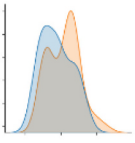
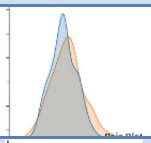
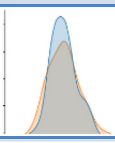
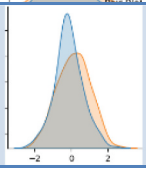
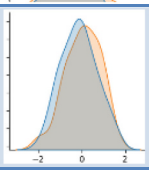
3.4 Dimensionality Reduction on the Dataset

PCA is employed to reduce the dimensionality of the dataset after extracting 29 features. The process starts with standardizing the data to have zero mean and unit variance. The covariance matrix is then computed to assess variable correlations, followed by the calculation of eigenvalues and eigenvectors. Eigenvectors define the principal component directions, while eigenvalues indicate their variance. Components with the largest eigenvalues are selected for projection, resulting in a lower-dimensional representation of the data. The PCA results and graphic distribution are summarized in Tables 1 and 2.

Table 1. Dimensionality Reduction Results

PC	Eigenvalues	Cumulative variance	Explained variance
PC 1	6.69826519	0.232	0.23182491
PC 2	5.08080807	0.408	0.17584522
PC 3	4.21769304	0.554	0.14597307
PC 4	2.23239542	0.631	0.07726252
PC 5	1.90579493	0.697	0.06595898
PC 6	1.6290246	0.753	0.05638004
PC 7	1.22280806	0.795	0.04232101
PC 8	1.18573757	0.836	0.04103802
PC 9	0.93145304	0.868	0.03223731
PC10	0.74576327	0.894	0.02581064
PC11	0.70306182	0.918	0.02433276

Table 2. Histogram Distribution Results of Train and Test Data Based on PCA Principal Components.

Principle Component	Data Train	Data Test
PC 1		
PC 2		
...	...	...
PC 10		
PC 11		

3.5 Classification Using SVM

The feature extraction results from facial images, post-PCA, are classified using a Support Vector Classifier (SVC) model. The dataset of 476 samples is divided into training (80%) and test (20%) sets. The training data is utilized to develop the classification model, while the test data evaluates its performance. The classification employs a one-vs-one binary approach, with four kernel types tested: radial basis function (RBF), polynomial, linear, and sigmoid. Each kernel's parameters include: RBF (C and gamma), polynomial (C, gamma, degree, and coef0), linear (C), and sigmoid (C, gamma, kernel, and coef0). The RBF kernel, with C and gamma values of 1000 and 0.1, respectively, yields the highest performance. The accuracy, sensitivity, specificity, and precision are 95%, 95%, 94%, and 94%, respectively. Figure 3 shows the confusion matrix, where 93.75% of normal images (45 out of 48) and 95.74% of stroke images (45 out of 47) are correctly classified.

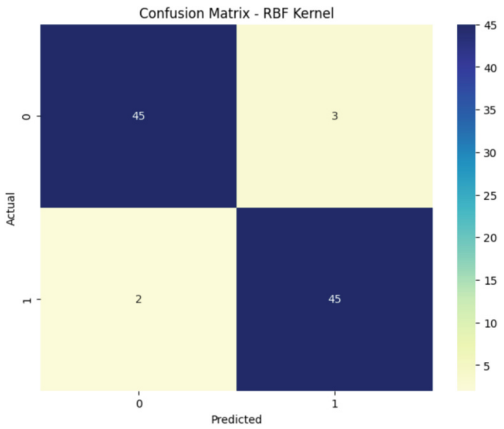


Fig. 3. Where class 0 represents normal and class 1 represents stroke.

4 Discussion

The proposed system, integrating Feedback Cascade Regression for facial alignment and SVM for classification, achieved a high accuracy of 95% in detecting facial weakness in stroke patients. This result demonstrates the effectiveness of advanced pre-processing techniques combined with robust classification models in improving detection accuracy.

4.1 Preprocessing of Facial Images

The preprocessing phase standardizes facial images for machine learning by applying grayscale conversion, resizing, rescaling, normalization, rotation, and facial landmark alignment. Initially, images are resized to 128x128 pixels while maintaining aspect ratios to reduce distortion and aid feature extraction. The resizing process also involves cropping to focus on the face, removing background, and ensuring that the resulting images are uniform. Rescaling adjusts these resized images to ensure the face is centered within the target dimensions, with borders added if necessary to preserve proportions. Image rotation corrects any tilt to standardize facial orientation across the dataset, enhancing consistency for landmark detection. Finally, facial landmark alignment uses a feedback cascade regression method, which iteratively refines landmark positions in three stages—salient, inner, and all—to improve detection accuracy. This comprehensive preprocessing ensures that facial images are prepared in a consistent manner, facilitating effective machine learning analysis.

4.2 Principal Component Analysis

After extracting 29 features, dimensionality reduction is performed using Principal Component Analysis (PCA) to address data errors. PCA results show that PC1, with an eigenvalue of 6.698, explains 23.2% of the total variance, while PC2, with an eigenvalue of 5.081, accounts for an additional 17.6%, making the cumulative variance explained by these two components approximately 40.8%. Overall, the first 11 principal components

capture around 91.8% of the data variance, indicating that most of the dataset's information is retained with these components. The data distribution, illustrated in Table 2, shows that PC1's histogram for the training data exhibits some overlap between the two classes, with class blue having a slightly higher peak. This overlap suggests that PC1 is not highly effective in class separation. PC2, however, shows a better separation between classes, with distinct peaks for class blue and class orange. For the test data, PC1 displays a similar distribution pattern as the training data, with significant overlap between classes. PC2 in the test data also demonstrates improved class separation, mirroring the trends observed in the training data, thus proving to be more effective in distinguishing the classes.

4.3 Hyperparameter SVM Classification

In the kernel and hyperparameter testing, the Radial Basis Function (RBF) kernel performed best with a parameter C of 100 and gamma of 0.1, achieving an accuracy of 95%. This combination, balancing a smaller margin with a lower gamma, effectively separates classes while reducing overfitting risks. The polynomial kernel showed optimal results with a C of 10 and degree of 3, yielding a 90% accuracy. This setup successfully maps data to a higher-dimensional space, improving separation. Conversely, the linear kernel achieved its highest accuracy of 75% with a C of 1000, but struggled with complex data patterns. The sigmoid kernel had varied results, with the best accuracy of 72% at a C of 0.1 and coef0 of 0.1, indicating limited effectiveness for this dataset. Overall, the RBF kernel with a C of 100 and gamma of 0.1 provided the most reliable performance.

4.4 Comparison with Related Studies

Compared to previous studies, such as the work by Zhuang et al. (2020) which achieved 94% accuracy using a combination of facial landmarks and Histogram of Oriented Gradients (HoG) features, the proposed method offers a slight improvement in accuracy. Additionally, the use of Feedback Cascade Regression enhances the model's ability to adapt to varying facial conditions, making it more robust against diverse patient expressions.This comparison is detailed in Table 3, which provides a summary of the performance metrics relative to other studies. This comparison is detailed in Table 3, which provides a summary of the performance metrics relative to other studies.

Table 3. Comparison with Previous Research

Research	Dataset	Class	Accuracy	Sensitivity	Specificity	Precision
Proposed method	Toronto	2	95%	95%	95%	94%
Landmarks + HoG (Zhuang et al., 2020) [7]		2	94%	–	–	–

4.5 Weaknesses of the Study

Despite the strong performance of the proposed model, certain limitations are evident. The study is confined to detecting facial weakness from still images, which could be expanded to video-based detection for a more dynamic assessment. While the method is computationally efficient and proficient at handling complex facial expressions, it remains reliant on image-based analysis. This limitation implies that the model may not fully capture the nuances of facial weakness observable in video sequences, where facial expressions and movements are more varied. Future research should focus on extending these capabilities to video analysis to enhance the model's applicability and robustness in real-world scenarios.

4.6 Implications of the Study

The findings of this study have significant implications for facial weakness detection in stroke patients. By achieving high accuracy and balanced performance metrics, the proposed SVM model with an RBF kernel demonstrates its potential as a reliable tool for early detection and diagnosis of facial weakness, crucial for timely intervention and treatment. The success of this model underscores the effectiveness of integrating advanced preprocessing techniques with machine learning methods, contributing valuable insights to clinical practice. Moreover, the study highlights the importance of employing comprehensive datasets and robust evaluation methods to improve model reliability and generalizability, setting a benchmark for future research in this field.

5 Conclusion

This study proposes a system for detecting and classifying facial weakness in stroke patients using facial images and Support Vector Machines (SVM) for classification. The primary goal is to evaluate the effectiveness of this system in distinguishing between facial weakness due to stroke and normal facial conditions. The dataset used in this study is the Toronto Neuro dataset, which includes labeled images of facial weakness. The dataset comprises 473 images, divided into two classes: 237 images of facial weakness and 236 normal images. Image preprocessing and feature extraction were performed, followed by classification using SVM. The highest accuracy achieved was 95%, with the optimal parameters being a C value of 100 and a gamma of 0.1. This result was obtained from test data, which was used to evaluate the model's performance on unseen data. Despite this high accuracy, there were some misclassifications, with instances where normal images were incorrectly classified as stroke and vice versa. This issue highlights the impact of varying facial expressions on feature extraction and classification accuracy. The study's strength lies in its advanced face alignment using feedback cascade regression, which enhances the system's adaptability to extreme patient conditions and complex facial expressions. Future research should focus on expanding the study to include video-based detection to capture the dynamics of facial expressions better, which may enhance classification accuracy. Additionally, further improvements in landmark detection methods, such as integrating more advanced feedback cascade regression techniques, could help address the current limitations related to extreme patient conditions.

Further investigations into the influence of image acquisition conditions on classification outcomes are also recommended to improve the model's robustness and generalizability.

References

1. Singh PK (2021) World Stroke Day. WHO South-East Asia. https://www.who.int/southeastasia/news/detail/28-10-2021-world-stroke-day, last accessed 2024/07/24
2. Schimmel M, Ono T, Lam OLT, Müller F (2017) Oro-facial impairment in stroke patients. J Oral Rehabil 44:313–326. https://doi.org/10.1111/joor.12486
3. Glober N et al (2016) Acute Stroke: Current Evidence-based Recommendations for Prehospital Care. West J Emerg Med 17:104–128. https://doi.org/10.5811/westjem.2015.12.28995
4. Nor AM et al (2004) Agreement between ambulance paramedic- and physician-recorded neurological signs with face arm speech test (FAST) in acute stroke patients. Stroke 35:1355–1359. https://doi.org/10.1161/01.STR.0000128529.63156.c5
5. Anatomy: Stroke vs Bell's Palsy, https://www.ebmconsult.com/articles/anatomy-stroke-vs-bells-palsy. Last accessed 24 07 2024
6. Rahajeng ACD, Nuh M, Hikmah NF (2020) An evaluation performance of kernel on support vector machine to classify the skin tumors in Dermoscopy image. In: 2020 International conference on computer engineering, network, and intelligent multimedia (CENIM). IEEE, pp 76–81. https://doi.org/10.1109/CENIM51130.2020.9297941
7. Hikmah NF, Hajjanto AD, Surbakti AF, Prakosa NA, Asmaria T, Sardjono TA (2024) Brain tumor detection using a MobileNetV2-SSD model with modified feature pyramid network levels. Int J Electr Comput Eng 14:3995. https://doi.org/10.11591/ijece.v14i4.pp3995-4004
8. Anggraini S, Sardjono TA, Hikmah NF (2024) Optimizing hyperparameters in multiview convolutional neural network for improved breast cancer detection in mammograms. TELKOMNIKA (Telecommunication Comput Electron Control 22:921. https://doi.org/10.12928/telkomnika.v22i4.25846
9. Zhuang Y et al (2020) Facial weakness analysis and quantification of static images. IEEE J Biomed Heal Informatics 24:2260–2267. https://doi.org/10.1109/JBHI.2020.2964520
10. Hao Y, Zhu H, Shao Z, Ma L (2019) Feedback cascade regression model for face alignment. IET Comput Vis 13:632–639. https://doi.org/10.1049/iet-cvi.2018.5347
11. Parra-Dominguez GS, Sanchez-Yanez RE, Garcia-Capulin CH (2021) Facial paralysis detection on images using key point analysis. Appl Sci 11:2435. https://doi.org/10.3390/app11052435
12. Paul L, Suman A (2012) Face recognition using principal component analysis method. Int J Adv Res Comput Eng Technol 1:932–936. https://www.researchgate.net/publication/318362885_Face_recognition_using_principal_component_analysis_method
13. Hsu C, Lin C (2002) A comparison of methods for multiclass support vector machines. IEEE Trans Neural Networks 13:415–425. https://doi.org/10.1109/72.991427
14. Yadav S, Shukla S (2016) 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. https://doi.org/10.1109/IACC.2016.25
15. Bandini A et al (2021) A new dataset for facial motion analysis in individuals with neurological disorders. IEEE J Biomed Heal Inf 25:1111–1119. https://doi.org/10.1109/JBHI.2020.3019242
16. Krig S (2014) Image pre-processing. In: Computer vision metrics. Apress, Berkeley, CA, pp 39–83. https://doi.org/10.1007/978-1-4302-5930-5_2

17. Kim H, Lee J (2019) Facial emotion recognition using convolutional neural networks. IEEE Trans Affect Comput 2:324–336
18. Saponara S, Elhanashi A (2022) Impact of image resizing on deep learning detectors for training time and model performance. Presented at the. https://doi.org/10.1007/978-3-030-95498-7_2
19. Vasanthi M, Seetharaman K (2022) Facial image recognition for biometric authentication systems using a combination of geometrical feature points and low-level visual features. J King Saud Univ - Comput Inf Sci 34:4109–4121. https://doi.org/10.1016/j.jksuci.2020.11.028
20. Guarin DL et al (2020) Toward an automatic system for computer-aided assessment in facial palsy. Facial Plast Surg Aesthetic Med 22:42–49. https://doi.org/10.1089/fpsam.2019.29000.gua
21. Nguyen HT, Tran MT (2021) Advances in facial feature detection using machine learning algorithms. Neurocomputing, pp 174–188
22. Juhong A, Pintavirooj C (2017) Face recognition based on facial landmark detection. In: 2017 10th biomedical engineering international conference (BMEiCON). IEEE, pp 1–4. https://doi.org/10.1109/BMEiCON.2017.8229173
23. Sun L, Dai J, Shen X (2021) Facial emotion recognition based on LDA and Facial Landmark Detection. In: 2021 2nd international conference on artificial intelligence and education (ICAIE). IEEE, pp 64–67. https://doi.org/10.1109/ICAIE53562.2021.00020
24. Kaewmahanin W, Rassameecharoenchai T, Jutharee W, Tongskulroongruang T, Wiphunawat P, Jennawasin T, Kaewkamnerdpong B (2022) Automatic facial asymmetry analysis for elderly stroke detection by using cosine similarity. In: 2022 19th international conference on electrical engineering/electronics, computer, telecommunications and information technology (ECTI-CON). IEEE, pp 1–4. https://doi.org/10.1109/ECTI-CON54298.2022.9795508
25. Greene JJ et al (2020) The spectrum of facial palsy: The MEEI facial palsy photo and video standard set. Laryngoscope 130:32–37. https://doi.org/10.1002/lary.27986
26. Kim H, Kim S, Kim Y, Park K (2015) A smartphone-based automatic diagnosis system for facial nerve palsy. Sensors 15:26756–26768. https://doi.org/10.3390/s151026756
27. Xu Y, Yang L (2019) Efficient facial recognition using improved SVM algorithms. IEEE Trans Neural Netw Learn Syst 4:1322–1335
28. Kremic E, Subasi A (2016) Performance of random forest and SVM in face recognition. Int Arab J Inf Technol 13:287–293
29. Parashar D, Agrawal DK (2020) Automated classification of glaucoma stages using flexible analytic wavelet transform from retinal fundus images. IEEE Sens J 20:12885–12894. https://doi.org/10.1109/JSEN.2020.3001972
30. Guo F, Li W, Tang J, Zou B, Fan Z (2020) Automated glaucoma screening method based on image segmentation and feature extraction. Med Biol Eng Comput 58:2567–2586. https://doi.org/10.1007/s11517-020-02237-2

Multiclass Glaucoma Classification in Fundus Retina Images Using Modified U-Net Segmentation and VGG16

Andini Vira Salsabilla Zulkarnain Pakaya(✉), Tri Arief Sardjono, Nada Fitrieyatul Hikmah, and Norma Hermawan

Institut Teknologi Sepuluh Nopember, Arief Rahman Hakim, 60111 Surabaya, Indonesia
andinivirap@gmail.com

Abstract. Glaucoma, a major cause of blindness, results from damage to the optic nerve due to increased intraocular pressure. Traditionally classified into normal or glaucoma categories, recent advancements have refined severity levels into mild (CDR 0.6–0.8), moderate (CDR 0.8–0.9), and severe (CDR > 0.9). However, manual Cup-to-Disc Ratio (CDR) calculations are expensive, require experienced clinicians, and often identify optic nerve damage only at advanced stages. This study aims to develop an automated system using Convolutional Neural Networks (CNN) to classify glaucoma severity into four detailed levels: normal, mild, moderate, and severe. The proposed method overcomes the limitations of manual CDR calculation by introducing a machine learning-based approach that provides a more detailed and accurate classification of glaucoma severity. Specifically, the study demonstrates the application of CNNs, with a focus on the VGG16 model architecture, to enhance glaucoma diagnosis. Utilizing the ORIGA dataset, the study involves pre-processing, optic disc and cup segmentation, feature extraction, and classification using the VGG16 model. The model's performance was assessed through key metrics, including accuracy, precision, recall, specificity, and F1 Score. The developed model achieved an accuracy of 88.36%, precision of 91.23%, recall of 87.82%, specificity of 97.41%, and an F1 Score of 87.41%, indicating its effectiveness in detecting and classifying glaucoma based on fundus retinal images. This automated classification system presents a reliable and efficient alternative to traditional diagnostic methods, potentially enabling earlier detection and more effective treatment planning for glaucoma patients. Implementing this automated system in clinical settings could significantly enhance the early diagnosis and management of glaucoma, thereby reducing the risk of vision loss in patients.

Keywords: Glaucoma · Fundus image · Glaucoma classification · Cup-to-disc Ratio · CNN

1 Introduction

Glaucoma, a group of eye diseases that progressively damages the optic nerve and leads to vision loss, is frequently associated with increased intraocular pressure (IOP) [1]. Managing glaucoma, particularly in its early stages, presents challenges due to

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 250–262, 2026.
https://doi.org/10.1007/978-981-95-9233-3_19

high fluid leakage rates [2]. Common detection methods include Intraocular Pressure Measurement (IPM), Function-Based Visual Field Tests (FBVFT), and Optic Nerve Head (ONH) examinations, which analyze metrics like the Cup-to-Disc Ratio (CDR) and Edge-to-Disc Ratio (EDR) [2, 3]. However, recent advancements in image processing and deep learning have improved glaucoma detection accuracy [4, 5]. For instance, CNN-based architectures such as U-Net have been successfully utilized for optic cup and disc segmentation to derive CDR [6], while the Gray-Level Co-occurrence Matrix (GLCM) has improved texture feature extraction for classification [7]. Despite these advancements, most studies have focused on binary or three-class classification models, limiting early diagnosis and prevention [8].

In this context, deep learning models like VGG16 have significantly enhanced classification accuracy by allowing for detailed categorization of glaucoma into four classes: normal, mild, moderate, and severe. This study aims to address the limitations of traditional diagnostic methods, showing how VGG16 can outperform previous models in accuracy and efficiency, providing a more robust automated tool for early glaucoma detection. Ahmad [1] explored clinical approaches that rely heavily on expert input, which offers valuable insight into combining clinical expertise with machine learning for more nuanced analysis. However, this approach's disadvantage is its reliance on clinician involvement, which may hinder scalability. Similarly, Bernabé et al. [2] and Civit-Masot et al. [3] used CNN-based techniques to classify eye diseases, implementing dual machine-learning systems that enhance segmentation precision but require high computational resources, which can limit their practicality in certain clinical settings. Parashar and Agrawal [4] employed SVM for glaucoma classification, successfully reducing false positives, yet they encountered difficulties with multi-class categorization, especially in differentiating mild and moderate glaucoma stages. Singh and Garg [5] applied fuzzy C-means clustering, which showed promising results in minimizing false positives but also struggled with complex multi-class classifications. By incorporating these advancements, this study develops a refined classification model that addresses both the technological and clinical complexities in glaucoma diagnosis.

2 Materials and Method

The study used the ORIGA dataset of 650 retinal images from the Singapore Eye Research Institute, with ground truth annotations for Cup-to-Disc Ratio (CDR), optic disc, and optic cup segmentation. Of these, 168 images are glaucomatous, and 482 are healthy. Preprocessing involved contrast normalization with CLAHE, median filtering, resizing to 256 $\times$ 256 pixels, green channel extraction for contrast enhancement, and cropping to focus on the optic disc and cup regions.

2.1 General System Design

Based on Fig. 1 The system starts by acquiring a retinal fundus image, which undergoes meticulous pre-processing to enhance quality. The green channel is extracted for its superior contrast in highlighting retinal structures. A median filter smooths the image

while preserving critical edges, and CLAHE enhances contrast. Color and shape normalization standardizes the images, followed by cropping to isolate the region of interest and sharpening to enhance detail.

The pre-processed image is input into a U-Net model for biomedical image segmentation, accurately identifying the optic cup and disc. These segmentation results enable the calculation of the Cup-to-Disc Ratio (CDR), a crucial metric for assessing glaucoma severity. Based on the CDR, retinal images are classified into severity categories: Normal, Mild, Moderate, or Severe, aiding clinical decisions. To further refine classification, texture features are extracted using Gray-Level Co-occurrence Matrix (GLCM) analysis and fed into a Support Vector Machine (SVM). Advanced CNN architectures like VGG16 are also employed to classify pre-processed images directly, leveraging deep learning techniques to recognize complex patterns and ensure precise severity classification across different cases of glaucoma.

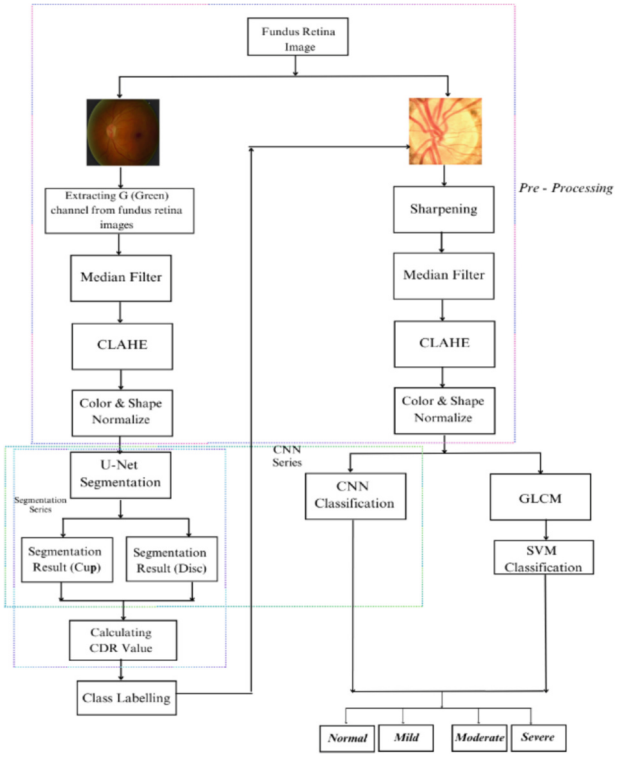


Fig. 1. Diagram System

2.2 ORIGA Dataset

This study utilized the open-source dataset ORIGA, which contains 650 retinal images annotated by trained professionals from the Singapore Eye Research Institute [9]. A

comprehensive set of image markers essential for the diagnosis of glaucoma is described. The system will be continually updated with more clinical truth images. The dataset used in this study focuses on glaucoma and consists of RGB images with a resolution of 3072 by 2048 pixels. It is available online as referenced in the source. The dataset is divided into two classes: normal and glaucomatous. It includes comprehensive ground truth information, such as the cup-to-disc ratio (CDR) values, segmentation ground truth of the optic disc (OD) and optic cup (OC), and classification ground truth for both normal and glaucomatous classes [10]. There are 650 images in the dataset, with 168 images classified as glaucomatous and 482 images as healthy [11].

2.3 Pre-processing Image

The pre-processing for fundus images is divided into two categories: original-sized fundus retina and cropped fundus retina. The pre-processing stage in this software design involves several key steps. Initially, it includes shape normalization and resizing to align with the original images. For fundus images, the pre-processing includes shape normalization to eliminate unnecessary areas [12], resizing to a standardized size of 256×256, color normalization to maintain consistent color across the images, applying Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance essential image details [13], using median filtering to reduce noise [14], and sharpening to improve the clarity of the fundus retina images [14].

Shape and color normalization adjust RGB values to ensure consistency across devices, resizing images to 256×256 pixels for efficient processing and CNN compatibility. Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to the green channel of fundus images to enhance contrast and improve detail visibility. Sharpening techniques like Unsharp Masking and High-pass Filtering are used to emphasize edges and details in the images. The median filter reduces noise while preserving important details, resulting in clearer representations of key structures such as the optic disc and cup, which are crucial for accurate classification.

2.4 U-Net Segmentation

The optic disc segmentation stage employs a pre-trained U-Net model [6]. This model is trained to accurately segment the optic disc in retinal images. Once the optic disc is segmented, its boundaries are identified and highlighted in the image. Subsequently, the optic cup segmentation stage begins. The U-Net model, which has been fine-tuned for optic cup segmentation, is utilized.

Image segmentation divides an image into segments based on specific criteria to facilitate understanding and analysis. In retinal fundus images, segmentation isolates blood vessels, the optic cup, and the optic disc, aiding in accurate measurement. The areas of the optic disc and optic cup are calculated using the segmented boundaries. Eq. (1) provides critical measurements for assessing the Cup-to-Disc Ratio (CDR) and identifying potential glaucoma cases [15].

$$\text{CDR Value} = \frac{\textit{optic cup diameter}}{\textit{optic disc diameter}} \tag{1}$$

where Optic Cup Diameter: Indicates the width of the central hollow region within the optic disc, Optic Disc Diameter: Indicates the total width of the optic disc, including both the optic cup and the surrounding rim, and CDR Value: Indicates the proportion of the optic disc occupied by the optic cup. Higher values can suggest glaucoma.

2.5 GLCM Extraction Features, Features Selection, and SVM

The GLCM (Gray-Level Co-occurrence Matrix) method transforms an image into numerical values based on texture characteristics, using a matrix with rows and columns at specific gray levels [16]. Feature extraction using the Gray Level Co-occurrence Matrix (GLCM) improves the model's adaptability to variations in fundus images by focusing on parameters like Contrast, Dissimilarity, Homogeneity, and Energy. This method considers the relationship between the frequency of two pixels with specific intensities. Additionally, it can analyze the optic disc, which has a finer texture compared to other parts of the retina [17]. The Pearson Correlation Function selects features correlated with the class [18]. Support Vector Machine (SVM) is a newly developed technique for solving various classification and regression problems [19]. A Support Vector Machine (SVM) classifies data by finding an optimal hyperplane, using kernels such as linear, RBF, sigmoid, and polynomial, with parameters C and gamma fine-tuned for better accuracy [20]. SVM can handle both linear and non-linear classification through kernel tricks

2.6 VGG 16 CNN Classification

VGG16 is a convolutional neural network (CNN) architecture introduced by the Visual Geometry Group (VGG) at the University of Oxford [21]. Many image preprocessing projects, such as classification and diagnosis, can be accomplished using deep learning models [22]. This CNN architecture is tailored for detecting glaucoma in retinal fundus images, featuring adaptable layers for enhanced performance. It includes input, convolutional, pooling, and fully connected layers [23]. The FC layers are trained for the final classification after multiple alternating convolutional and pooling layers. CNN has been frequently used in image classification tasks with great effectiveness [24]. VGG16, developed by the Visual Geometry Group at Oxford, is renowned for its simplicity and high performance in image recognition tasks. It consists of 16 layers, including 13 convolutional and 3 fully connected layers. Each convolutional layer uses 3×3 filters to extract detailed features, while 2×2 pooling layers reduce spatial dimensions to optimize computation. VGG16 is widely adopted in image recognition and computer vision applications.

There are four operations in general of CNN consisting of Conv layer, RelU Activation, Pooling Layer, and Fully Connected Layer operations [25]. The operation for Convolutional Layers **Purpose to** Extract features from the input is in Eq. (2).

$$\mathrm{S(i,j)} = \mathrm{a} = \sum_{a=0}^{f-1} \sum_{a=0}^{f-1} I(\mathrm{i} + \alpha, \mathrm{j} + \mathrm{b}) \cdot \mathrm{K}\ (\alpha, \mathrm{b}) \tag{2}$$

where is the input, K is the filter, and is the feature map. The operation for ReLU Activation to Introduce non-linearity is in Eq. (3).

$$f(x) = \max(0, x) \tag{3}$$

where x is the input to the activation function. Eq. (4) represents the operation for Pooling Layer to downsample the feature map.

$$Y(i, j) = \max(X(i + \alpha, j + b)) \tag{4}$$

where X is the input and Y is the pooled output. The operation for the Fully Connected Layer to Classify the features is shown in Eq. (5).

$$Z = W \cdot X + b \tag{5}$$

where W is the weight matrix, X is the input, and Z is the output.

3 Result

This section will elaborate on all the results of the experiments, such as reprocessing, segmentation, GLCM Features, SVM, and CNN.

3.1 Pre-Processing Image

The research starts with pre-processing to enhance image quality using the ORIGA dataset, augmented to 1001 images to address class imbalance (originally 650 images: 168 Glaucoma, 482 Non-Glaucoma). Pre-processing involves shape normalization, cropping, and resizing images to 256 × 256 pixels, focusing on relevant areas like the optic disk. CLAHE is applied to each RGB channel to enhance contrast, highlighting details like blood vessels and lesions. Sharpening techniques, such as Unsharp Masking and High-pass Filtering, emphasize edges and important details. The median filter reduces noise by replacing pixel values with the median of neighboring pixels, preserving critical details in both original and cropped images.

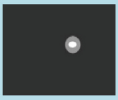
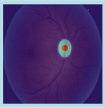
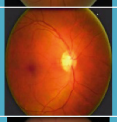
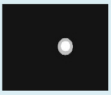
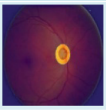
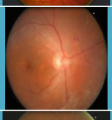
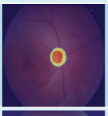
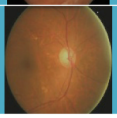
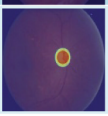
3.2 Segmentation U Net

The ORIGA original image dataset was used to train a U-Net model for segmenting Cup and Disc areas to calculate the Cup to Disc Ratio (CDR). The U-Net, implemented with PyTorch and MONAI, used a modified ResNet34 encoder. Input images (256 × 256 × 3 RGB) and ground truth images (256 × 256 grayscale) were processed through the encoder, middle convolution layers, and upsampled by the decoder using skip connections. Key hyperparameters were: 1 epoch, batch size of 3, and learning rate of 0.0001, with Cross Entropy loss and Adam optimizer. The training showed low loss, indicating good performance, with final segmentation reaching a validation accuracy of 96.81%.

Five sample images with ground truths in Table 1 were evaluated for segmentation accuracy, with results overlaid on the originals. CDR values for 650 images were calculated, saved as CSV, and labeled with one-hot encoding. Segmentation closely matched

the ground truths, indicating success. Overlaid images visualized accuracy, and CDR values aided glaucoma diagnosis. The ORIGA dataset showed class imbalance (33 normal, 120 mild, 323 moderate, 80 severe), which was addressed with data augmentation (Table 1).

Table 1. Comparison CLAHE Grid Tile

Image Original	Segmentation Result	*Overlay*	CDR	*Class*
			0.301	Normal
			0.444	*Mild*
			0.583	*Moderate*
			0.711	*Severe*

3.3 GLCM Extraction Features, Features Selection, and SVM

After segmentation, texture features were extracted using GLCM to compute contrast, dissimilarity, correlation, energy, and homogeneity. This process starts by selecting pixel pairs at a certain distance and angle, building a matrix based on the frequency of specific gray levels. These metrics aid in texture recognition and classification. Support Vector Machine (SVM) classification was then applied using various kernels: linear, polynomial, RBF, and sigmoid. The dataset was split into 80% training, 10% validation, and 10% testing. Hyperparameter tuning and cross-validation determined the optimal parameters. From Table 2 we know that The RBF kernel achieved the highest accuracy at about 80% with specific C and gamma combinations. The polynomial kernel performed well with lower degrees, while the sigmoid kernel was less effective.

Table 2. GLCM Features

Image 201					
Class	*Contrast*	*Dissimilarity*	*Correlation*	*Energy*	*Homogeneity*
Normal	30.081	3.74994	0.98551	0.0286	0.30079
Image 047					

(*continued*)

Table 2. (*continued*)

Image 201					
Mild	34.074	3.70748	0.98940	0.0597	0.30079
Image 118					
Moderate	2.2381	0.96146	0.99826	0.0526	0.62844
Image 158					
Severe	46.617	4.66548	0.99277	0.0389	0.276558

3.4 CNN VGG16 Classification

The study evaluated the classification accuracy of a multi-class glaucoma detection model using the VGG16 architecture pre-trained on ImageNet, tested on the ORIGA dataset. Various hyperparameters such as epochs (75, 100, 200), batch size (16, 32), and learning rates (0.1, 0.01, 0.001) were tested and shown in Fig. 2. Adam with a 0.001 learning rate proved to be the most effective, providing high accuracy and low loss across different epochs, while RMSProp also showed strong performance at higher epochs and lower learning rates. Best Optimizer and Learning Rate: Adam with a 0.001 learning rate is the most effective for training the VGG16 model on the ORIGA dataset, providing high accuracy and low loss across different epochs. RMSProp also showed strong performance, especially at higher epochs and lower learning rates. As shown in Fig. 2, the training accuracy achieves 95%, and the validation accuracy stabilizes around 88.36%, despite some fluctuations. The training loss decreases to 20%, while the validation loss remains at 50%, indicating adequate model performance.

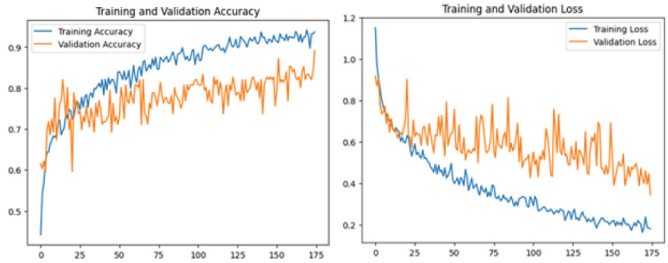


Fig. 2. (a) and (b) Training and Validation Accuracy and Loss Graphic

The confusion matrix in Fig. 3 provides a detailed assessment of the model's performance across four glaucoma classes: Normal, Mild, Moderate, and Severe. For the Normal class, the model successfully classified 35 instances, yielding an accuracy of 89.7%, but it misclassified 4 instances as Moderate, resulting in a misclassification rate of 10.3%. In the Mild class, the model excelled with perfect accuracy, correctly classifying all 39 instances (100%) with no misclassifications.

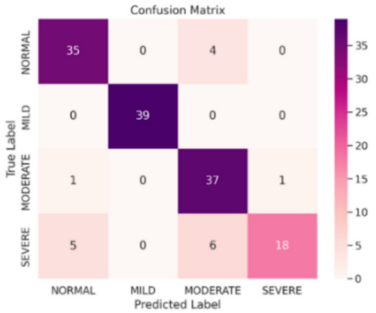


Fig. 3. Confusion Matrix Highest Accuracy

For Moderate glaucoma, the model accurately classified 37 out of 39 instances, achieving an accuracy of 94.9%. However, it misclassified 1 instance as Normal (2.6%) and another as Severe (2.6%), indicating some confusion between these classes. In the Severe category, the model classified 18 instances correctly, resulting in a 75% accuracy rate. Nonetheless, it misclassified 6 instances as Moderate (25%) and 5 as Normal (20.8%), highlighting challenges in differentiating Severe cases from the other classes. Overall, while the model demonstrates strong performance, particularly in the Mild and Moderate classes, it struggles with accurately identifying Severe cases, suggesting a need for further refinement. This is evident despite the model exhibiting good learning capabilities, as indicated by decreasing training loss and increasing training accuracy throughout the training process.

4 Discussion

Pre-processing begins by standardizing the ORIGA dataset to JPG format. Images are cropped to focus on the optic disc and resized to 256 × 256 pixels. Color normalization balances light intensity and color across RGB channels, and CLAHE enhances contrast, particularly in the green channel. Sharpening techniques like Unsharp Masking and High-pass Filtering improve image sharpness, while a median filter reduces noise. Segmentation focuses on the optic cup and disc areas, validated by matching masks to ground truth. The CDR values for 650 images are calculated and stored in a CSV file, categorized using one-hot encoding. Overlay visualizations confirm segmentation accuracy, with the U-Net model achieving 96.81% accuracy and a loss of 0.00043.

To address the class imbalance in the ORIGA dataset, image augmentation techniques such as rotation and contrast adjustment were used, resulting in a balanced and diverse dataset, enhancing the model's generalization and performance in classifying retinal conditions [26]. GLCM feature extraction highlighted texture differences across glaucoma severity levels, with the "Severe" class showing the highest contrast and dissimilarity, and the "Moderate" class the highest homogeneity. SVM classification, using an 80:20 data split, tested RBF, polynomial, and sigmoid kernels. The RBF kernel, with $C = 1000$ and gamma $= 0.1$, achieved the highest accuracy at 80%. The polynomial kernel reached 76.67% accuracy, while the sigmoid kernel had the lowest at 70%. The RBF kernel's superior handling of nonlinearity made it the best choice for multiclass

glaucoma classification. The VGG16 model processes 224 × 224 RGB images through five convolutional blocks with MaxPooling2D layers, followed by GlobalAveragePooling2D, a dense layer with 1024 neurons, and a 4-neuron output layer for classification. Hyperparameter tuning involved adjusting the number of convolutional layers, epochs, learning rate, optimizer, and batch size. Increasing epochs (75, 100, 200) improved accuracy and reduced loss, with the best results at 200 epochs (accuracy: 0.8700). Among the optimizers tested (Adam, SGD, RMSprop), Adam provided the best performance with stable, high accuracy, and F1-Score, followed by RMSprop, while SGD showed less stable and lower performance.

Table 3. Comparison with previous research

Methods	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1 Score (%)	Class
Propose Method	88.36	91.23	87.82	97.41	87.4	4
(Jibhakate et al., 2022) [24]	74. 00	–	–	–	–	2
GFI-C Deep Learning Martins J, et al., 2020) [27]	76.57	–	–	80. 00	–	2
VGG Capsnet (Singh LK et al., 2022) [28]	88.52	82.00	82.00	82.00	86.0	2
VGG16 (Simonyan and Zisserman, 2015) [29]	88.22	–	–	–	–	2
LS – SVM (Kumar Dheeraj et al., 2020) [17]	92.21	81.45	94. 04	92.53	87.3	3
GBDT (Tang J, et al., 2020) [30]	84.30	–	–	79.00	–	2

The comparative analysis of various glaucoma prediction methods highlights the strengths of each approach. Table 3 shows that the proposed method demonstrates robust

performance with an accuracy of 88.36%, precision of 91.23%, recall of 87.82%, specificity of 97.41%, and an F1 Score of 87.41% for 4 classes, indicating exceptional effectiveness, particularly in precision and specificity. In contrast, VGG16 (Jobakat et al., 2022) shows lower accuracy (74.00%) and lacks comprehensive metrics. GFI-C Deep Learning (Martins J et al., 2020) offers moderate performance with 76.57% accuracy and 80.00% specificity but lacks precision and recall data. VGG Capsnet (Singh LK et al., 2022) achieves 88.52% accuracy with balanced metrics but slightly lower specificity. VGG16 (Simonyan and Zisserman, 2015) and GBDT (Tang J et al., 2020) provide strong accuracy but lack detailed metrics. LS-SVM (Kumar Dheeraj et al., 2020) excels with 92.21% accuracy and strong recall but lower precision. The proposed method stands out for balanced, high-performing metrics, making it superior for effective glaucoma prediction.

Enhancing image pre-processing, such as adding filters and using high-quality datasets, significantly impacts CNN classification accuracy. Improving segmentation through better decoder architectures or using pre-trained models and evaluating metrics like IoU and Dice coefficient is crucial. Simplifying models and parameters reduces computational load and considers hardware specifications. CNN classification accuracy can be improved by adding layers or adjusting hyperparameters. Using balanced, multi-class fundus image datasets further enhances the development. Although image augmentation helped, the inherent class imbalance may still affect the model's ability to detect rarer conditions. Additionally, the high computational load of the complex CNN model, with its deep layers and extensive tuning, may not be practical for all users. Implication from this method's high precision and specificity make it promising for early glaucoma detection, potentially improving diagnostic tools and mass screening programs. The study also highlights the importance of advanced image processing techniques like CLAHE and color normalization, which enhance data quality and could be refined for other medical imaging tasks. Additionally, effective hyperparameter tuning and kernel selection are crucial for optimizing model performance, underscoring the need for continued research in machine learning for medical diagnostics. The model exhibits strong accuracy, precision, and specificity, making it reliable for predictions across different classes. Effective pre-processing and high segmentation accuracy, particularly using the U-Net model, significantly contributed to its performance. The study also provides a thorough comparative analysis, underscoring the method's strengths relative to others. However, the focus on optic disc and cup segmentation may overlook other critical features, and the high computational demands could limit its use in resource-constrained environments. Additionally, the lack of external validation raises concerns about its generalizability to other datasets, potentially limiting clinical applicability.

5 Conclusion

This study aimed to develop an effective multiclass classification system for glaucoma detection, categorizing cases into normal, mild, moderate, and severe stages. The proposed method achieved this through a comprehensive process involving pre-processing, U-Net-based segmentation, feature extraction, and selection, and classification using both SVM and CNN models. Quantitatively, the study reported an accuracy of 88.36%

for the CNN classification, with a sensitivity of 86.67% and a specificity of 96.06%. The U-Net segmentation process reached an accuracy of 96.81%, while the SVM classifier demonstrated a specificity of 93.09% and a precision of 79.17%. Future work should focus on expanding the dataset to improve model generalization, incorporating additional features beyond the optic disc and cup for enhanced accuracy, and validating the model on external datasets to ensure its clinical applicability. Additionally, optimizing the computational efficiency of the model could make it more accessible for widespread use.

References

1. Ahmad SS (2018) Glaucoma suspects: a practical approach. https://doi.org/10.4103/tjo.tjo_106_17
2. Bernabe O, Acevedo E, Acevedo A, Carreno R, Gomez S (2021) Classification of eye diseases in fundus images. IEEE Access 9:101267–101276. https://doi.org/10.1109/ACCESS.2021.3094649
3. Civit-Masot J, Dominguez-Morales MJ, Vicente-Diaz S, Civit A (2020) Dual machine-learning system to aid glaucoma diagnosis using disc and cup feature extraction. IEEE Access 8:127519–127529. https://doi.org/10.1109/ACCESS.2020.3008539
4. Parashar D, Agrawal D (2020) Automated classification of glaucoma using retinal fundus images. In: 2020 1st IEEE international conference on measurement, instrumentation, control and automation, ICMICA 2020. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICMICA48462.2020.9242702
5. Singh LK, Khanna M, Thawkar S, Singh R (2022) Collaboration of features optimization techniques for the effective diagnosis of glaucoma in retinal fundus images. Adv Eng Softw 173. https://doi.org/10.1016/j.advengsoft.2022.103283
6. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK (2018) Medical image analysis using convolutional neural networks: a review. https://doi.org/10.1007/s10916-018-1088-1
7. Patil, J., Chaudhari, S.: Determination of glaucoma grade with cup to disc ratio. (2021).
8. Alvionita V, Nuh M, Hikmah NF (2020) Data balancing techniques evaluation on convolutional neural network to classify the diabetic retinopathy of fundus image. In: CENIM 2020—Proceeding: international conference on computer engineering, network, and intelligent multimedia 2020. Institute of Electrical and Electronics Engineers Inc, pp 354–359. https://doi.org/10.1109/CENIM51130.2020.9297940
9. Rakshitarora: Mahotas: Median Filter.
10. Wedianto A, Sari HL, Yanolanda Suzantri H (2016) Analisa Perbandingan Metode Filter Gaussian, Mean Dan Median Terhadap Reduksi Noise, vol 12
11. Zhang Z et al (2010) ORIGA: an online retinal fundus image database for glaucoma analysis and research. In: 2010 annual international conference of the IEEE engineering in medicine and biology. IEEE, pp 3065–3068. https://doi.org/10.1109/IEMBS.2010.5626137
12. Kumar VVNS, Harinath Reddy G, Giri Prasad MN (2023) A novel glaucoma detection model using Unet++-based segmentation and ResNet with GRU-based optimized deep learning. Biomed Signal Process Control 86. https://doi.org/10.1016/j.bspc.2023.105069
13. Anggraini S, Sardjono TA, Hikmah NF (2024) Optimizing hyperparameters in multiview convolutional neural network for improved breast cancer detection in mammograms. TELKOMNIKA (Telecommunication Comput Electronics Control) 22:921. https://doi.org/10.12928/telkomnika.v22i4.25846

14. Firdi NP, Sardjono TA, Hikmah NF (2022) Using pectoral muscle removers in mammographic image process to improve accuracy in breast cancer. J Biomimetics Biomaterials Biomed Eng 55:131–142. https://doi.org/10.4028/p-35cy9o
15. Kumar JRH, Seelamantula CS, Kamath YS, Jampala R (2019) Rim-to-disc ratio outperforms cup-to-disc ratio for glaucoma prescreening. Sci Rep 9. https://doi.org/10.1038/s41598-019-43385-2
16. Paramastri GA, Pramunendar RA (2021) Glaucoma detection through fundus images using radial basis function with color and GLCM feature extraction. In: 2021 international seminar on application for technology of information and communication (iSemantic). IEEE, pp 222–227. https://doi.org/10.1109/iSemantic52711.2021.9573220
17. Priyanka Kumar D (2020) Feature extraction and selection of kidney ultrasound images using GLCM and PCA. Proc Comput Sci 167:1722–1731. https://doi.org/10.1016/j.procs.2020.03.382
18. Almusallam N, Tari Z, Chan J, Fahad A, Alabdulatif A, Al-Naeem M (2021) Towards an unsupervised feature selection method for effective dynamic features. IEEE Access 9:77149–77163. https://doi.org/10.1109/ACCESS.2021.3082755
19. Chan K, Lee T-W, Sample PA, Goldbaum MH, Weinreb RN, Sejnowski TJ (2002) Comparison of machine learning and traditional classifiers in glaucoma diagnosis. IEEE Trans Biomed Eng 49:963–974. https://doi.org/10.1109/TBME.2002.802012
20. Vedaldi, A., Zisserman, A. 2016 VGG convolutional neural networks practical.
21. Rahajeng ACD, Nuh M, Hikmah NF (2020) An evaluation performance of kernel on support vector machine to classify the skin tumors in dermoscopy image. In: 2020 international conference on computer engineering, network, and intelligent multimedia (CENIM). IEEE, pp 76–81. https://doi.org/10.1109/CENIM51130.2020.9297941
22. Raghavendra U, Bhandary SV, Gudigar A, Acharya UR (2018) Novel expert system for glaucoma identification using non-parametric spatial envelope energy spectrum with fundus images. Biocybern Biomed Eng 38:170–180. https://doi.org/10.1016/j.bbe.2017.11.002
23. Skouta A, Elmoufidi A, Jai-Andaloussi S, Ouchetto O (2022) Hemorrhage semantic segmentation in fundus images for the diagnosis of diabetic retinopathy by using a convolutional neural network. J Big Data 9:78. https://doi.org/10.1186/s40537-022-00632-0
24. Jibhakate P, Gole S, Yeskar P, Rangwani N, Vyas A, Dhote K (2022) Early glaucoma detection using machine learning algorithms of VGG-16 and Resnet-50. In: 2022 IEEE region 10 symposium (TENSYMP). IEEE, pp 1–5. https://doi.org/10.1109/TENSYMP54529.2022.9864471
25. Velpula VK, Sharma LD (2023) Multi-stage glaucoma classification using pre-trained convolutional neural networks and voting-based classifier fusion. Front Physiol 14. https://doi.org/10.3389/fphys.2023.1175881
26. Parashar D, Agrawal DK (2020) Automated classification of glaucoma stages using flexible analytic wavelet transform from retinal fundus images. IEEE Sensors J 20:12885–12894. https://doi.org/10.1109/JSEN.2020.3001972
27. Martins J, Cardoso JS, Soares F (2020) Offline computer-aided diagnosis for glaucoma detection using fundus images targeted at mobile devices. Comput Methods Prog Biomed 192. https://doi.org/10.1016/j.cmpb.2020.105341
28. Singh LK, Khanna M, Pooja (2022) A novel multimodality based dual fusion integrated approach for efficient and early prediction of glaucoma. Biomed Signal Process Control 73:103468. https://doi.org/10.1016/j.bspc.2021.103468
29. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. (2014).
30. Guo F, Li W, Tang J, Zou B, Fan Z (2020) Automated glaucoma screening method based on image segmentation and feature extraction. https://doi.org/10.1007/s11517-020-02237-2/Published

Utilizing IoT-Based Breath Analysis for Early Detection and Monitoring of Diabetes: A Deep Learning Approach

Nilakshi Maruti Mule[1(✉)] and Dipti Durgesh Patil[1,2]

[1] STES's Smt Kashibai Navale College of Engineering, Savitribai Phule Pune University, Pune, India
nilmule@gmail.com
[2] MKSSS's Cummins College of Engineering for Women, Pune, India

Abstract. Volatile organic compounds (VOCs) are helpful biomarkers because they can reveal various diseases' underlying pathophysiology and physical abnormalities. Analyzing exhaled breath is a simple and non-invasive way to monitor metabolic rate. Diabetes is a complex metabolic disease. Oxidative stress, inflammatory syndrome, diabetes, and hypertension are all intricately linked. This research utilizes an IoT-based breath analyzer to collect real-time data on diabetes and healthy persons. The study aims to design and develop its IoT-based breath analyzer device to identify the exhaled breath gases of diabetes persons. This method was used to test a group of 26 people with diabetes and 17 people without the disease. The data is gathered through the Android module and serially acquired through the USB cable. The Cool-Term open-source software is used for data acquisition. After the real-time data collection, the disease characteristics are assessed by a deep learning system. The model successfully distinguishes between diabetes and non-diabetes exhaled breath samples with a 97% success rate.

Keywords: Diabetes · VOCs · Human Exhaled Breath · Bi-LSTM · Hypertension

1 Introduction

Diabetes Mellitus (DM) is a long-term metabolic problem that happens when the body doesn't have enough insulin. It can be told apart by having abnormal blood sugar levels, changed metabolism of proteins, carbohydrates, and fats, and a higher risk of vascular problems [1, 2]. This condition's potential origin could be impaired insulin action or inadequate insulin synthesis, resulting in a syndrome characterized by principally lipid, protein, and sugar metabolic abnormalities [3]. Insulin, a hormone, converts carbohydrates, sugar, and other nutrients into energy [4]. Hyperglycemia is an indication of diabetes and is brought on by insufficient insulin. Many parts of the body can get worse because of this condition, such as the eye, the renal glomerulus, peripheral neuropathy, microvascular and macrovascular coronary artery disease, and stroke [5]. Unmanaged diabetes leads to various phases of ketoacidosis and hyperosmolar coma. While the

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 263–278, 2026.
https://doi.org/10.1007/978-981-95-9233-3_20

exact cause is uncertain, it appears that the same factors that contribute to ketoacidosis, specifically those that lead to dehydration, also have a role in triggering this illness [6]. Long-term diabetes in a patient leads to foot ulcers, impaired kidney function, and retinopathy [7]. Diabetes and hypertension frequently coexist, leading to an accelerated advancement of kidney deterioration [8].

The World Health Organization (WHO) says that diabetes is a problem that is getting worse quickly in India. About 8.7% of people aged 20 to 70 have diabetes. The International Diabetes Federation (IDF) says that there are 90 million people with diabetes in Southeast Asia and 537 million people with diabetes around the world in 2021. India is one of the seven countries in the IDF sea region. But by 2045, 151.5 million more people will live in this region's countries.

Type 1 (T1DM) and Type 2 DM (T2DM) are the two most common forms of the disease. Type 1 diabetes typically manifests throughout childhood, and its severity increases with advancing age. The condition is insulin-dependent diabetes mellitus [9, 10]. Type 1 diabetes is the targeted death of pancreatic beta islet cells by cell-mediated autoimmunity. It causes a complete lack of insulin and raises the chances of developing diabetic ketoacidosis (DKA) [11]. Individuals diagnosed with type 2 diabetes are commonly referred to as having mature-onset diabetes or non-insulin-dependent DM. Approximately 90% of all cases of diabetes are this type, and most people get it later in life. Two physiological problems are what make T2DM happen: insulin resistance and not producing enough insulin. Even though T2DM is more common than T1DM and has a more complicated case, little is known about how it starts [11]. It is typically characterized by gradually increasing blood sugar levels and lacks apparent symptoms. Elevated insulin resistance and reduced insulin secretion. This results in a relative insulin deficiency, leading to elevated glucose levels in the blood [12].

The majority of people who develop type 1 diabetes do so in their early thirties or early forties. These symptoms are distinct from those seen in type 2 diabetes. Symptoms can appear suddenly. Patients often experience symptoms such as frequent urination, constant thirst, increased appetite, fatigue, and weight loss. Diabetic ketoacidosis (DKA) is frequently the first sign of type 1 diabetes. Frequent manifestations include abdominal pain, dry skin, diminished foot sensation, rapid and deep respiration, and vomiting. Most people with type 2 diabetes are identified as having symptoms in their middle years or later, often when they turn 40. The onset of type 2 diabetes mellitus symptoms is gradual and subtle. Upon physical examination, patients who are diagnosed with glycosuria or hyperglycemia typically do not manifest any observable symptoms. Fatigue, sluggishness, pruritus, skin infections, visual impairments, mood fluctuations, headaches, dizziness, and muscle cramps are the primary signs and symptoms [13].

Although blood tests can identify prediabetes and diabetes, having them checked in a licensed laboratory is advisable to guarantee the correctness and dependability of findings. Nevertheless, they can be employed as a rapid signal of elevated blood glucose levels. The efficacy of glucose measuring equipment already in use is insufficient for accurate diagnosis. Testing enables the identification and management of prediabetes and diabetes before the onset of complications [14]. Polyuria, polydipsia, lethargy, loss of weight, blurred vision, increased hunger, and diabetic dermatomes are the primary signs of diabetes. Any laboratory test that detects diabetes without obvious symptoms must

be double-checked with another measurement [15]. Due to the discomfort experienced by most patients during blood tests, there is a requirement for a non-painful method to monitor T1DM and T2DM [16, 17].

People breathe in oxygen and breathe out carbon dioxide and other VOCs. The many VOCs released when people breathe result from the biological processes inside our bodies. The significant things that make up human breath are nitrogen, oxygen, carbon dioxide, water, and others [18]. Augment the quantity of acetone and carbon dioxide expelled during respiration in individuals with diabetes. Medical research often involves analyzing human breath, which offers a non-invasive and quick way to screen and diagnose complex conditions like cancers and acute infections [19]. In addition, this technology provides several benefits when compared to conventional diagnostic methods. These include painless procedures and sampling that do not require the expertise of professional technicians [20].

Chunhui Deng et al. [4] developed a novel technique for detecting acetone in human breath. This method combines gas chromatography-mass spectrometry (GC-MS) and solid-phase microextraction (SPME) with on-fiber derivation. Collecting breath gas samples is a common practice for studying healthy individuals and those with diabetes. Keeping track of small levels of acetone in individuals' breath could be an extra method for detecting diabetes. This approach utilizes a straightforward, efficient, sensitive, and environmentally friendly technology. In their study, M. Righettoni et al. [16] examined the exhaled breath of participants using a portable sensor specifically developed to identify acetone. The sensor was created using nanostructured sheets of silicon-doped tungsten trioxide (WO3) that were produced through a flame-based process. Using Tedlar bags is necessary for collecting breath samples. Proton Transfer Reaction Time-of-Flight Mass Spectrometry (PTR-MS) is commonly employed to assess and quantify glucose levels in an individual's blood.

Ke Yan et al. [20] developed innovative breath analysis equipment that utilizes specialized chemical sensors to identify biomarkers in exhaled breath. Assessing the device's efficacy involves evaluating breath samples from 295 healthy individuals and 279 individuals with diabetes. The expected BGL has an average absolute relative error of 21.7%. The diabetes screening test has a sensitivity of 91.51% and a specificity of 90.77%. The primary benefit of this research is the monitoring of blood glucose variability in diabetes. The breath is collected in a Tedlar bag for analysis; prolonged duration may lead to degradation of breath concentration. In their study, Guo et al. [21] present a breath analysis system that detects acetone in an individual's breath using chemical sensors. This novel system can assist in the detection of diabetes and the surveillance of glycemic levels. The study's strength gives direct breath measurements, but no control arm maintains plasma glucose at baseline levels.

In their investigation, Clair Turner et al. [22] examined the exhaled airways of eight individuals diagnosed with type 1 diabetes using Selective Ion Flow Tube Mass Spectrometry (SIFT-MS). Patients undergo "Insulin Clamp" examinations, during which insulin and glucose are administered to reduce their blood glucose levels from average to low progressively. Breath acetone levels in individuals with type 1 diabetes can fluctuate in response to changes in blood sugar levels and metabolic conditions, according to studies. This study analyzed the exercise condition based on real-time VOCs, but the

system did not diagnose any disease, and the sample size was smaller. PTR-MS was used in a study by King et al. [23] to assess breath data and important hemodynamic and respiratory characteristics from a carefully thought-out experiment. The study's strength is that it tracked breath CO_2 dynamics over a long time and simultaneously took repeated measurements of breath CO_2 and blood glucose and plasma insulin levels. The study can't show that the method works as a diagnostic tool that can be used instead of or in addition to current methods or to get more people, like overweight kids, to go for diabetes screenings.

Furthermore, a methodology for integrating SPME/GC-MS data with VOC profiles produced by PTR-MS is presented in this document. The correlation between respiratory patterns, cardiac output, and blood pressure, as well as their influence on the measured concentration of breath, is investigated in this study. They assist in identifying distinct VOCs that indicate transitions from activity to rest due to ventilation or blood flow alterations.

A study by Lichar Dillon et al. [24] examined individuals with standard glucose tolerance, prediabetes, and early-stage diabetes to analyze their excretion of breath $13CO_2$. The participants were subjected to a 13C-labeled oral glucose tolerance test (OGTT) to identify potential variations. These findings suggest that this approach could be beneficial for individuals at risk of prediabetes or early-stage diabetes, allowing them to assess their condition without invasive blood tests. This study's advantage is that breath samples and room air were taken simultaneously. This was done so that ambient gas concentrations could be deducted from the breath samples and displayed as net values. The method's limited dataset makes evaluating its applicability in other demographics difficult. Ten healthy people were studied by Pietro R. Galassetti et al. [25] by collecting samples of insulin, blood glucose, ambient air, and exhaled breath gas. These samples were collected in the same manner as any other oral glucose tolerance test. An analysis was conducted on the gas sample using GC-MS. The calculation of glucose profiles demonstrates a significant correlation with actual glucose readings, with an average individual correlation of 0.70. In addition, all subjects had a correlation coefficient greater than or equal to 0.41. The study employed multiple regression analysis, with glucose, ethanol, and acetone as independent variables. This study's strength lies in examining the relationships between breath acetone and blood glucose, A1C, and other metrics through clinical testing on Type 1 diabetes patients with a ringdown breath acetone analyzer. The relationship between type 2 diabetes patients and breath acetone was not thoroughly established; however, other compounds in human breath may also serve as biomarkers for diabetes.

Using a prototype cavity ring-down breath acetone analyzer, Chuji Wang et al. [26] performed clinical trials on people diagnosed with diabetes. The researchers thoroughly examined the correlations between breath acetone and many parameters, such as the type of diabetes onset, gender, age, weight, blood glucose level, and A1C value. The study revealed that 16 individuals diagnosed with type 1 diabetes had an average breath acetone concentration of 2.19 ppm by volume (ppmv), substantially more than the 0.48 ppmv observed in the 15 healthy individuals. The sample size is significantly less. A study conducted by Gang-Ting Fan et al. [27] has shown that the Hadamard transform-gas chromatography/mass spectrometry (HT-GC/MS) approach is practical in identifying

small quantities of acetone in human breath samples using gas chromatography (GC). The researchers identified a spectrum of respiratory acetone concentrations among the eight subjects, varying from 0.1 to 1 ppmv. In contrast, the values of the four diabetic patients ranged between 1 and 10 ppm. The advantage of this work is that lower operating temperatures require less power, small response, and recovery time, but the dataset used is less.

Marco Righettoni et al. [28] successfully created a portable acetone sensor. The sensor is made from nanostructured sheets of Si-doped epsilon-wo3, which are created using a flame and then annealed in place. By fine-tuning the gas flow speeds and sensor temperatures, they achieved an impressive detection limit of 20 parts per billion (ppb) for acetone. The response time was very efficient, ranging from 10 to 15 s, while the recovery time was slightly longer, ranging from 35 to 70 s. The study's strength is real-time breath ethylene and ammonia monitoring with type 2 diabetes using a CO_2 LPAS system, but it required to expand the applicability of breath analysis in clinical diagnosis.

The proposed method identifies diabetes by utilizing a set of gas sensors. The objective of this study is to ascertain whether the examination of exhaled breath can distinguish between individuals with hypertension diabetes and those in healthy persons. The existing systems used for exhaled breath analysis are Gas Chromatography-Mass Spectrometry (GC-MS) and Electronic Nose (E-nose). The GC-MS is a costly, heavy instrument and requires skilled operators. E-nose is cost-effective but mostly used in fragrance companies. The use of E-nose systems in the medical field till not proved [29]. This study employs the TGS-822, TGS-2600, TGS-2602, MQ-7, and MQ-9 gas sensors. These sensors are compact, cost-effective, suitable for automated measurement, and can be used for real-time and uninterrupted detection. In addition, they possess a broad spectrum of dynamic range. The IoT-based analysis device does not require any skilled operators. It can be used at home, in rural areas, and in laboratories to diagnose early disease.

1.1 Research Objectives:

- The objective is to identify the specific gas that diabetic persons exhale concerning their blood pressure. The purpose is to propose an efficient approach for collecting breath samples.
- The objective is to collect real-time data through the IoT-based breath analyzer.
- The study investigates the disparities in exhaled breath composition between individuals with diabetes and those without health conditions.
- The goal is to evaluate the efficiency of the IoT-based breath analyzer device using a bidirectional long-short-term memory (BiLSTM) algorithm.

1.2 Major Contributions:

- To develop a breathalyzer that utilizes IoT technology.
- Gather samples from individuals with and without diabetes to detect immediate fluctuations in the gases exhaled from their breath.
- The objective is to utilize deep learning techniques to differentiate between the exhaled gasses of those with diabetes and those who are healthy.

2 Proposed Methodology

Researching and analyzing VOCs to apply sensor technologies for breath analysis, illness detection, and diagnosis is necessary. VOCs in the breath are a sign of sickness. This compound mixture is not disruptive to the device's sensors. Five commercially available semiconductor metal oxide gas sensors produced by Figaro USA, Inc. are utilized to create the sensor array. The employed gas sensors are TGS 822, TGS 2600, TGS 2602, MQ-7, and MQ-9. The specifications of the sensors are illustrated in Table 1. The sensing components of the Figaro gas sensors utilized in this study consist of a metal oxide.

The semiconductor layer exhibits minimal conductivity in pristine air. In the presence of observable gasses, conductivity can be transformed into an electrical output signal using a primary electrical circuit. All sensors employed in this work are cross-responsive sensors rather than gas-specific sensors. The sensors exhibit varying responses to identical gas mixtures, and the reaction of a single sensor differs when exposed to diverse gas mixtures. Rather than quantifying an individual chemical from exhaled breath, these sensors generate a voltage signal reflecting exhalation constituents' amalgamation [30].

The Design of portable measurement equipment incorporates five commercially available gas sensors. The properties shown in Table 1 are compiled using the datasheet for each sensor. The Design consists of five mechanical parts. The blower, sample chamber, air valve, non-return valve, and containment analyzer, also known as the sensor module, are the parts that make up the assembly, as shown in Fig. 1. Fig. 2 displays the electronic system responsible for analyzing the content.

Initially, the compressor is utilized to extract gas inlets from the individual. The gas, which has been expelled, proceeds to traverse the succeeding compartments. Automatically closing, the check valve prevents the backflow of air. An air reservoir within the sampling chamber retains the gas sample before the gas analyzer analyses. Once the gas analyzer has processed all the gas in the sampling chamber, the air valve is sealed. A content analyzer integrates different gas sensors into one chamber or multiple chambers, each with a separate collection of gas sensors. Their functionality determines the sensors' dependence. The gas analyzer is equipped with various gas sensors to determine the composition of the patient's sample.

A digital interface established a link between the sensors and the microcontroller unit. The data is acquired through Cool-Term using the serial port. Serial ports are interconnected to enable the smooth transmission of data and various types of information. It is beneficial to get data in a text file type and show it as either text or hexadecimal. The data samples obtained are shown in Fig. 3. The item is sent to the doctor for examination and to obtain the diagnosis.

Table 1. A brief description of the gas detector.

Sensors	Features	Gas Identification	Principle of Sensing	PPM Detection Range	Time to Preheat
TGS-822	This sensor exhibits remarkable stability and reliability over an extended duration, possesses a lengthy lifespan, is economically efficient, and showcases a heightened sensitivity to organic solvent vapors, specifically ethanol.	Organic Solvents, Volatile Vapors, Carbon Monoxide (CO)	Tin Dioxide Semiconductor (SnO_2)	300 -Ethanol	Over a week
TGS-2600	This product boasts several notable characteristics, including its deficient power consumption, outstanding sensitivity to gaseous air pollutants, an impressively extended lifespan, and attractively affordable pricing. A simple electrical circuit has a manageable and valuable design.	Hydrogen, CO	Metal Oxide Semiconductor (MOS)	1~30 Hydrogen	A week
TGS-2602	Increased sensitivity to toxic vapors and VOCs. This device possesses a robust capacity to identify gaseous air pollutants and demands less electrical power. Extended lifetime can be achieved using a simple electrical circuit with minimal dimensions.	Hydrogen, Ammonia, VOCs	MOS	1~30 -EtOH	A week
MQ-7	Hypersensitivity to carbon monoxide. Long-lasting and persistent existence.	CO	SnO_2	20–2000-CO	A minimum of 48 h
MQ-9	High vulnerability to long-lasting carbon monoxide, methane, and liquefied petroleum gas exposure.	CO	SnO_2	20–2000-CO, 500–10000-Methane, LPG	A minimum of 48 h

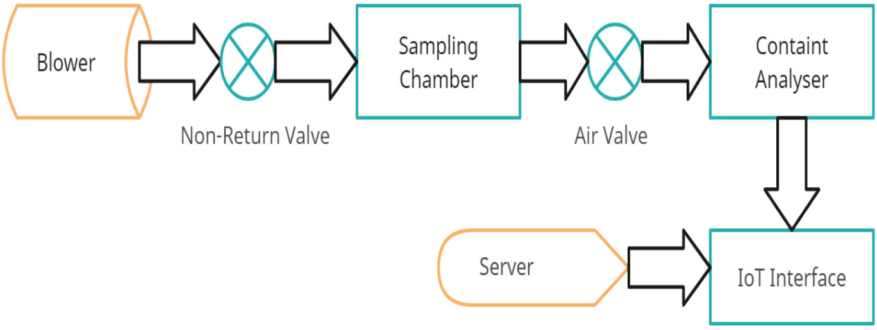


Fig. 1. The Fundamental Framework of The Prototype

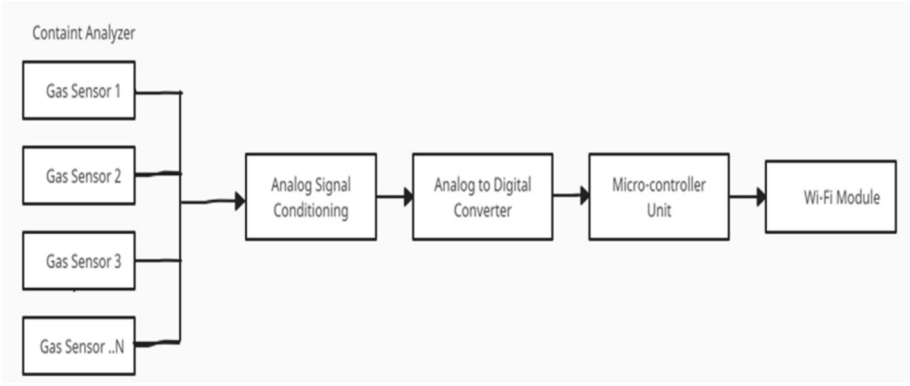


Fig. 2. Containment Analyzer Electronic System

The container analyzer has six sensors: DHT-11, TGS-822, TGS-2600, TGS-2602, MQ-3, and MQ-7. An appropriately positioned exhaust valve removes the sensor module's exhaled air. Air is extracted from the sensor module. When it's time to take a fresh breath sample, the device will display a signal.

2.1 Data Collection

The first step involves gathering real-time data through the IoT-based breath analyzer device. The tests were conducted with volunteer patients diagnosed with hypertension, diabetes, and healthy individuals as shown in Table 2, all of whom provided informed consent. During the sample collection process, no dietary restrictions are imposed. The dataset comprises 26 individuals diagnosed with diabetes. The dataset comprises various male and female individuals aged between 30 and 65. The study includes 17 healthy individuals aged between 25 and 50 years. The sampling stage was conducted with the setup of the device. Each device component can be disassembled rather than equipped with a sensor module. Upon completing the assembly of all device components, the sensors measure the initial air sample readings.

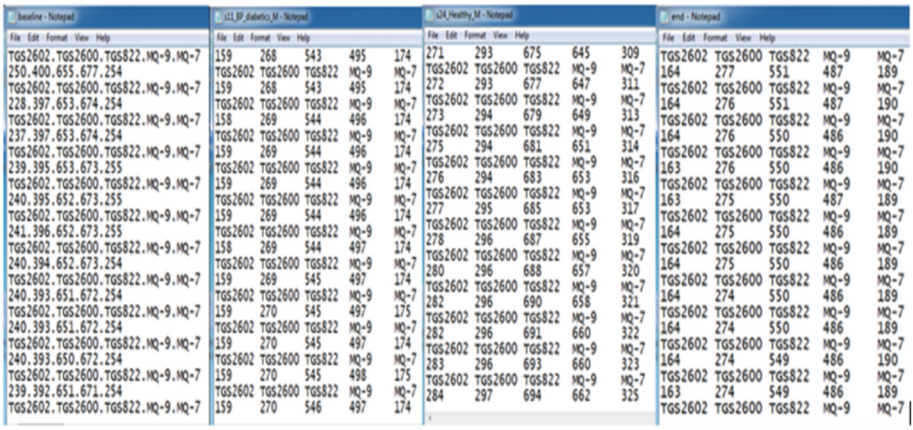


Fig. 3. Collected Data Samples

Table 2. Data Samples

Type	Number of Samples	Age	Female	Male
Diabetes	26	30–65	14	12
Healthy	17	25–50	08	09

2.2 Data Analysis

The deep learning method is utilized for data analysis once the data has been preprocessed. The MinMax scalar function is used for scaling purposes. The data normalization method preserves the original distribution's shape but shifts and rescales the values to fit within a specific range, usually 0 or 1; after data preprocessing, the deep learning method is applied for data analysis. Fig. 4 shows how the gas detectors responded to a breath sample taken from a diabetic patient with hypertension and a healthy human specimen. In the resulting graphs, the sensor analog value is indicated on the X-axis, measured in PPM, and the time step value is displayed on the Y-axis, measured in seconds.

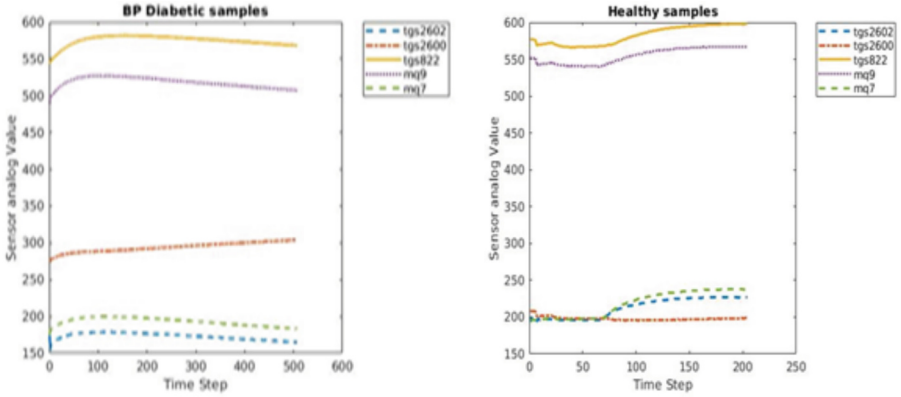


Fig. 4. The response of the gas sensors to patients with hypertension, diabetes, and healthy person samples

The graph analysis offers evidence of the disparities among the breath samples. Bi-LSTM algorithms are used in deep learning and artificial intelligence methods to examine the gathered data. Bi-LSTMs are mainly employed to learn and categorize sequential input. These networks detect persistent relationships between several points in a series of data. The data is classified into baseline measurements (before and post-breath sample readings) and end measurements (after-breath sample readings) for healthy individuals and individuals with diabetes. Afterwards, the sensor classifies the data. The rows in Fig. 4 and Fig. 5 represent the time steps, while the columns indicate the sensors' responses to each recorded breath sample regarding analog values.

The dataset is partitioned into two subsets for training and testing. The two parts are XTrain and YTrain: sequence training and segment. Twenty-five percent of the data is used for testing in the Bi-LSTM model, whereas 75 % is used for training. First, for each observation, the lengths of the sequences are declared. The variable Xtrain is utilized to denote the quantity of observations. After acquiring the sequence, its magnitude is ascertained through a comparison with the total number of observations. The duration of the sequence is either increased or decreased with each iteration of the loop. This research describes the Bi-LSTM network architecture. Set 5 as the input size and 60 as the hidden unit specification for the Bi-LSTM layer. The final piece of the sequence will be the output. To build various categories, add a densely linked layer of a size determined by the particular disease, followed by a softmax layer. List all of the available training programs. The solver should be designated as "Adam," the gradient threshold should be established at 1, and the values for "Maximum Epochs," "Mini-Batch Size," and "Maximum Epochs" should all be set to 500. The CPU is better suited for training because it can efficiently process brief sequences and mini-batches. The current value of the "Execution Environment" should be changed to "CPU." The LSTM network can be trained via the designated settings.

Integrate the data acquired from the collection of breath samples into the test results. Once the test set is loaded, the sequences can be classified into the Xtest and Ytest groups based on their properties. The Xtest system has a cell array containing sequences of different lengths, each with a size of 25. The Ytest vector is a vector that includes labels for the five sensors. The LSTM network is trained using mini-batches of sequences adjusted to have the same lengths. Sort the test results according to the training data's mini-batch size [31]. The device's effectiveness is assessed by collecting breath samples from diabetic patients and comparing them to healthy ones. The early-stage diagnosis of diabetes in these patients is established based on their clinical history. The breath samples are acquired utilizing the procedure described previously [32]. Fig. 4 illustrates the differentiation between samples that have diabetes and those that do not. Based on the graph, individuals with diabetes emit ethanol, acetone, carbon monoxide, methane, and isobutene through exhalation. The sensors provide crucial data on exhaled glucose. Fig. 5 demonstrates the system's ability to use its initial and final measurements. The lack of gas in the initial measurement resulted in all sensors registering baseline data, which is depicted as a linear trend on the graph.

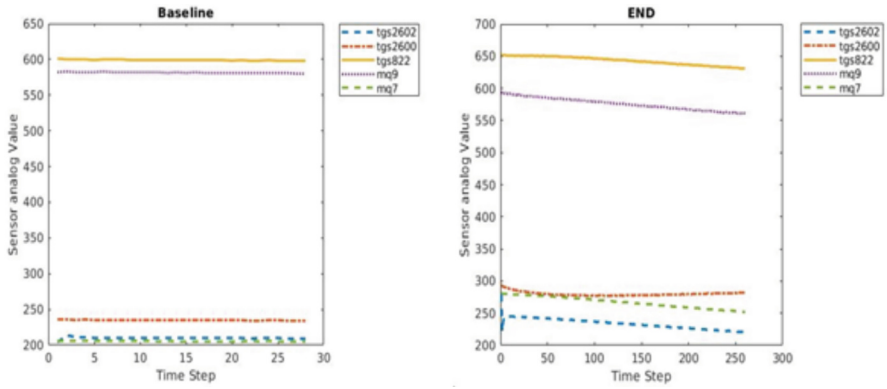


Fig. 5. Response of the gas sensor before the breath sample (baseline), the result, and breath end sample results

3 Results

The training procedure for the proposed deep learning is illustrated in Fig. 6. The figure displays the classification accuracy and training losses for each iteration. Initially, the loss is substantial, resulting in poor classification accuracy. However, as the number of repetitions rises, the training loss diminishes, enhancing accuracy.

Fig. 6. Model accuracy and loss of training and testing

The model was trained using an 80:20 ratio for the training and testing set. The results of the model-building phase showed that the BiLSTM algorithm achieved an accuracy of 97%, precision of 96%, recall of 99%, and the F1-score of 97%. Sensitivity is 98%, and specificity is 92%. They indicated an excellent ability to distinguish between the positive and negative classes, as shown in Fig. 7. The ROC curve also showed that the model had an AUC of 96%, as shown in Fig. 8.

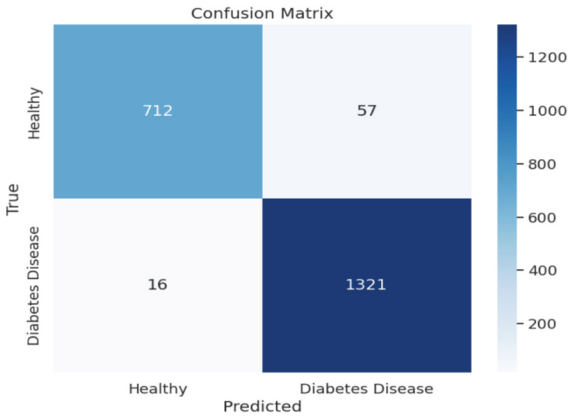


Fig. 7. Confusion Matrix

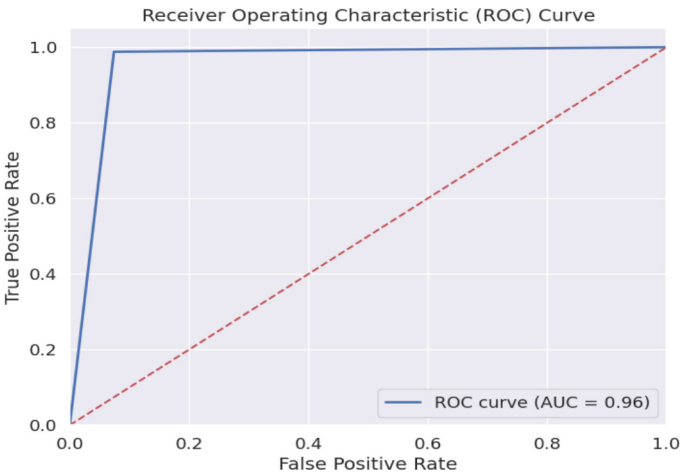


Fig. 8. ROC Curve

This study found that the sensors used to construct the sensor array captured distinct breath patterns in sick and healthy individuals, as reported previously. Additionally, it is observed that there is a variation in the respiratory cycle during the beginning and after the measurement cycle. This phenomenon can be attributed to the existence of both inhaled and exhaled gas components. The analysis of Fig. 4 reveals that the gas composition is unaffected by age and gender. The application of several sensor technologies in diabetes detection yielded significant and encouraging outcomes. The existing methods, namely GC-MS and E-nose, are expensive and challenging.

4 Discussion

This study presents an IoT-based breath analyzer device that incorporates an array of five gas sensors. TGS-822, TGS2600, TGS-2602, MQ-7, and MQ-9, along with a microcontroller board utilized and a deep learning algorithm created. This study illustrates that a sensor array combined with a classification algorithm can differentiate between diabetes patients and healthy individuals. Examining multiple samples from all participants validated this differentiation, indicating that the VOC profiles of exhaled breath exhibit notable disparities between healthy individuals and those with diabetes. This suggests that the breath analyzer device could serve as a valuable instrument for disease diagnosis shortly. This investigation employs an innovative method for determining the least number of sensors necessary to forecast disease conditions effectively. The pertinent studies concerning non-invasive diabetes detection are discussed in the related work section of the introduction. Numerous studies have employed exhaled breath samples for the diagnosis of diabetes.

Nevertheless, these investigations utilize different sensing modules. As a result, it would be unsuitable to directly compare the outcomes of the data classification methods analyzed in other studies. The limited sample size may challenge generalizing the findings and pursuing additional investigations on more extensive datasets. To address

the limitations of Bi-LSTM, it is essential to explore various alternative methods for predicting exhaled breath. The proposed method could facilitate the early detection and diagnosis of diabetes-related conditions. The investigation improved diagnostic accuracy, leading to elevated levels of classification precision. The proposed approach could simplify the breath collection process. Moreover, investigation and examination can broaden this method to identify other conditions, including pulmonary and respiratory diseases.

5 Conclusion

Utilizing five distinct MOS sensors, the equipment designed to differentiate between the breath patterns of diabetic patients and healthy controls achieved a 97% accuracy rate. This paper documents the observation of an IoT-based breath analyzer designed to detect acetone, isobutene, ethanol, carbon monoxide, and methane in the exhaled breath of individuals with diabetes. This study used the Bi-LSTM algorithm to assess the efficacy of the test by collecting and analyzing breath samples. This inexpensive, rudimentary IoT breath analyzer has shown superior performance in detecting diabetes. Further investigation of this approach is warranted for its application in diagnosing additional respiratory and pulmonary disorders such as COPD, asthma, and cancer. Trained with a large dataset, our model enables non-invasive early disease diagnosis and prognosis. The cost-effective, portable, and noninvasive diabetes detection and monitoring device described here can be efficiently employed in remote regions.

References

1. American Diabetes Association Diagnosis and classification of diabetes mellitus (2009) Diabetes Care, pp S62–S67
2. Saasa V, Malwela T, Beukes M, Mokgotho M, Liu C-P, Mwakikunga B (2018) Sensing technologies for detection of acetone in human breath for diabetes diagnosis and monitoring. Diagnostics 8:12. https://doi.org/10.3390/diagnostics8010012
3. Karalliedde J, Gnudi L (2014) Diabetes mellitus, a complex and heterogeneous disease, and the role of insulin resistance as a determinant of diabetic kidney disease. Nephrol Dial Transplant gfu405. https://doi.org/10.1093/ndt/gfu405
4. Deng C, Zhang J, Yu X, Zhang W, Zhang X (2004) Determination of acetone in human breath by gas chromatography–mass spectrometry and solid-phase microextraction with on-fiber derivatization. J Chromatogr B 810:269–275. https://doi.org/10.1016/j.jchromb.2004.08.013
5. Tripathi K (2008) Essentials of medical pharmacology. Jaypee Brothers Medical Publishers (P) Ltd. https://doi.org/10.5005/jp/books/10282
6. Walker, B.R. Colledge, N.R. 2008 Davidson's principles & practice of medicine
7. Rang HP, Dale MM, Ritter JM, Flower RJ, Henderson G. (2011) Rang and Dales pharmacology.
8. Pittas AG, Greenberg AS (2003) Contemporary diagnosis and management of diabetes. Assocs in Medical Marketing Co Inc
9. Report of the expert committee on the diagnosis and classification of diabetes mellitus (2003) Diabetes Care 26:s5–s20. https://doi.org/10.2337/diacare.26.2007.S5
10. Mohan H (2018) Textbook of pathology. Jaypee Brothers Medical Publishers

11. Wells BG et al (2008) Pharmacotherapy principles and practice. The McGraw Hill Companies
12. Surya S, Salam AD, Tomy DV, Carla B, Kumar RA, Sunil C (2014) Diabetes mellitus and medicinal plants-a review. Asian Pacific J Trop Dis 4:337–347. https://doi.org/10.1016/S2222-1808(14)60585-5
13. Park K (2009) Preventive and social medicine. M/S Banarsidas Bhanot
14. Cooke DW, Plotnick L (2008) Type 1 diabetes mellitus in pediatrics. Pediatr Rev 29:374–385. https://doi.org/10.1542/pir.29-11-374
15. Righettoni M, Schmid A, Amann A, Pratsinis SE (2013) Correlations between blood glucose and breath components from portable gas sensors and PTR-TOF-MS. J Breath Res 7:037110. https://doi.org/10.1088/1752-7155/7/3/037110
16. Minh TDC et al (2011) Noninvasive measurement of plasma glucose from exhaled breath in healthy and type 1 diabetic subjects. Am J Physiol Metab 300:E1166–E1175. https://doi.org/10.1152/ajpendo.00634.2010
17. Mazzone PJ (2008) Analysis of volatile organic compounds in the exhaled breath for the diagnosis of lung cancer. J Thorac Oncol 3:774–780. https://doi.org/10.1097/JTO.0b013e31817c7439
18. O'Hara ME, Clutton-Brock TH, Green S, Mayhew CA (2009) Endogenous volatile organic compounds in breath and blood of healthy volunteers: examining breath analysis as a surrogate for blood measurements. J Breath Res 3:027005. https://doi.org/10.1088/1752-7155/3/2/027005
19. Buszewski B, Kęsy M, Ligor T, Amann A (2007) Human exhaled air analytics: biomarkers of diseases. Biomed Chromatogr 21:553–566. https://doi.org/10.1002/bmc.835
20. Yan K, Zhang D, Wu D, Wei H, Lu G (2014) Design of a breath analysis system for diabetes screening and blood glucose level prediction. IEEE Trans Biomed Eng 61:2787–2795. https://doi.org/10.1109/TBME.2014.2329753
21. Guo D, Zhang D, Li N, Zhang L, Yang J (2010) Diabetes identification and classification by means of a breath analysis system. https://doi.org/10.1007/978-3-642-13923-9_6
22. Turner C, Walton C, Hoashi S, Evans M (2009) Breath acetone concentration decreases with blood glucose concentration in type I diabetes mellitus patients during hypoglycaemic clamps. J Breath Res 3:046004. https://doi.org/10.1088/1752-7155/3/4/046004
23. King J et al (2010) Dynamic profiles of volatile organic compounds in exhaled breath as determined by a coupled PTR-MS/GC-MS study. Physiol Meas 31:1169–1184. https://doi.org/10.1088/0967-3334/31/9/008
24. Dillon EL et al (2009) Novel noninvasive breath test method for screening individuals at risk for diabetes. Diabetes Care 32:430–435. https://doi.org/10.2337/dc08-1578
25. Galassetti PR et al (2005) Breath ethanol and acetone as indicators of serum glucose levels: an initial report. Diabetes Technol Ther 7:115–123. https://doi.org/10.1089/dia.2005.7.115
26. Wang C, Mbi A, Shepherd M (2010) A study on breath acetone in diabetic patients using a cavity Ringdown breath analyzer: exploring correlations of breath acetone with blood glucose and Glycohemoglobin A1C. IEEE Sensors J 10:54–63. https://doi.org/10.1109/JSEN.2009.2035730
27. Fan G-T, Yang C-L, Lin C-H, Chen C-C, Shih C-H (2014) Applications of Hadamard transform-gas chromatography/mass spectrometry to the detection of acetone in healthy human and diabetes mellitus patient breath. Talanta 120:386–390. https://doi.org/10.1016/j.talanta.2013.12.025
28. Righettoni M, Tricoli A, Gass S, Schmid A, Amann A, Pratsinis SE (2012) Breath acetone monitoring by portable Si:WO3 gas sensors. Anal Chim Acta 738:69–75. https://doi.org/10.1016/j.aca.2012.06.002
29. Mansouri S, Boulares S, Alhadidi T (2020) Non-invasive measurement of blood glucose by breath analysis. IEEJ Trans Electr Electron Eng 15:1457–1464. https://doi.org/10.1002/tee.23216

30. Binson VA, Subramoniam M (2021) Design and development of an e-nose system for the diagnosis of pulmonary diseases. Acta Bioeng Biomech 23:35–44
31. Meshram DA (2023) D.D.P.: predicting sour or sweet: exploring advance DL methods for odor perception based on molecular properties. Int J Intell Syst Appl Eng
32. Dharmale G, Patil D, Shekapure S, Chougule A (2024) AI based medicine intake tracker. In: Nature-inspired methods for smart healthcare systems and medical data. Springer Nature Switzerland, Cham, pp 25–38. https://doi.org/10.1007/978-3-031-45952-8_2

Development of a Smart Punching Bag for Punch Force Level Identification Using Fuzzy Logic

Lilik Anifah[1](✉), Nurhayati[1], Puput Wanarti Rusimamto[1], Muhamad Syarifuddien Zuhrie[1], and Haryanto[2]

[1] Department of Electrical Engineering, Universitas Negeri Surabaya, Ketintang, Surabaya, Indonesia
lilikanifah@unesa.ac.id

[2] Department of Electrical Engineering, Universitas Trunojoyo Madura, Bangkalan, Indonesia

Abstract. Exercise using a punching bag is very useful for the body, as long as it is done according to the individual's abilities and not excessively. So far, punching bags have not been equipped with a level identifier of the force of the punches performed. The purpose of this study is to propose a smart punching bag to identify the force level of the punches that have been performed using fuzzy. The method applied in this study is fuzzy mamdani, with triangle function membership, and defuzzification using centroid and bisector. The input of the system is the result of the FSR 402 sensor and accelerometer, while the output of the system is the level of the punches that have been performed. Based on the experiments that have been carried out, the proposed approach punching bag for punch force level identification in self-defense training using fuzzy logic can be implemented, although using different defuzzification parameters produces different output results. However, this difference is not too large, namely the average value of the delta difference between the two experiments is −0.01175, the minimum value is −0.388, and the maximum value is 0.3645. This research implies that a monitoring system approach is obtained from punches hitting the punching bag. The contribution of this research is expected to help the training process when doing punching bag hitting exercises. When doing a punch, later it can be known how much force hits the punching bag.

Keywords: Smart · Punching Bag · Power · Identification · Fuzzy Logic

1 Introduction

Physical activity is closely related to mental health and quality of life [1]. Exercise is one of the physical activities that is very beneficial for the body. The selection of sports and training must be wise according to a person's condition because each sport has specific and unique characteristics and must follow the interests and abilities of the exerciser [2]. The benefits of exercise are reducing the risk of coronary heart disease and all causes, of cancer, and cardiovascular death, in addition to reducing weight. Other benefits include

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 279–290, 2026.
https://doi.org/10.1007/978-981-95-9233-3_21

increasing blood lipids, fasting blood glucose, blood pressure, resting cardiovascular function, cardiorespiratory fitness, and bone strength [3]. Of course, this activity must be in accordance with a person's abilities and capacities because when compared to someone inactive, individuals who do excessive physical activity are feared to experience increased anxiety, stress levels, and depression [4]. One of the sports options is martial sports, some examples of martial sports are pencak silat, wushu, taekwondo, karate, judo, and boxing [5]. Martial sports have a positive impact on mental health, self-related constructs, internal and external emotion regulation, well-being, cognitive functioning, and certainly have a positive impact on body strength. It is undeniable that this sport also has the potential to affect anxiety, self-related characteristics, depression, attention, learning, memory, planning, and changes in brain structure [6]. Sports training can be done without equipment or with equipment, one of the equipment used is a punching bag. Boxing training using a punching bag for six weeks can significantly reduce BMI, especially in the last weeks [7].

Research on punching bags is currently very rare. Research on efforts to analyze punches has been conducted by several previous researchers. Introduction to the recognition and classification of martial sports punches, including recognition of punches in pencak silat using acceleration sensors and supervised learning [8]. Research using acceleration sensors is also used in recognizing punches in karate and combined with CNN [9]. The use of punch trackers has also been used to analyze punches in martial sports, and this technology has been able to track well [10]. Punch recognition using Support Vector Machine (SVM) and Random Forest classifiers has been carried out and can recognize punches with very satisfactory results [11] and using Neural Network has also been carried out [12].

The punching bag currently used is a punching bag that cannot identify the power and direction of the movement made by the athlete. This information is very much needed to determine the progress of the training that has been done. The punching bag technology that has been developed is a punching bag equipped with a microenergy generator that will mark when the punch hits the specified target [13]. The use of a gauge sensor to measure the power of the punch, but the results of this study state that the use of this type is not suitable because it is difficult to hit the center of the sensor when hit by the athlete [14]. Other research is a camera-based numerical analysis based on punch speed information [15].

The difference with previous research is that research [16, 17] is a research attempt to analyze the punch from its movement, without knowing how much force is used to hit. While this research attempts to analyze the punch movement style. This information is very useful for users to know about the strength of their punches. This research is to complement so that in addition to knowing the movement, information about the strength of the punch that has been done is also known.

Previous research has been able to calculate the force that hits the punching bag, but the drawback is that only one point area can measure the force. This research is expected to measure the punching force anywhere in the punching bag area so that users can be more free to hit the punching bag. The purpose of this study is to propose a smart punching bag for punch force level identification using fuzzy logic. This study can help punching bag users in analyzing the punches that have been done. Previously, if a punch

has been done, we cannot know the level of the punch that has been done. With this study, punching bag users can analyze the punches that have been done during training.

2 Methods

2.1 Stage of Research

This research stage is the design of hardware and software for embedded systems used on punching bags, data collection, embedded system analysis, designing and implementing fuzzy logic, and system testing stages. The sensors used here are FSR 402 and an accelerometer Sensor. The FSR 402 sensor measures the force exerted by the athlete when hitting the punching bag, while the accelerometer is used to measure acceleration and determine the direction of the punch that has been done by the athlete. Force Sensing Resistors (FSR) are two terminals made from a mixture of polymers and conductive nanoparticles [18]. This FSR sensor has been used in several studies, including to measure grip strength using the Linear Vector Machine (SVM) algorithm. Based on the experiments that have been carried out, this system has an accuracy of 90.8% [19]. It is also used for pipeline monitoring and oil spillage control [20, 21]. Based on the research above, this study is sure that FSR is the right one to measure the force of the punch that hits the punching bag. In addition, an accelerometer sensor is used which is useful for measuring acceleration and informing the direction of the punch. This sensor has been used in previous studies, namely identifying a person's activity from signals received by the accelerometer [22]. The accelerometer sensor is also used to investigate the mechanics of mild traumatic brain injury, the results of this study showed that it can calculate orientation and angular velocity as intermediate steps [23].

2.2 Stage of Research

The system design is depicted in Fig. 1, FSR 402 and the accelerometer is connected to the wemos. Furthermore, information about the FSR force and acceleration of the punching bag will be obtained, which will then become fuzzy input. The output of the system is the level of force of the punch that hits the punching bag.

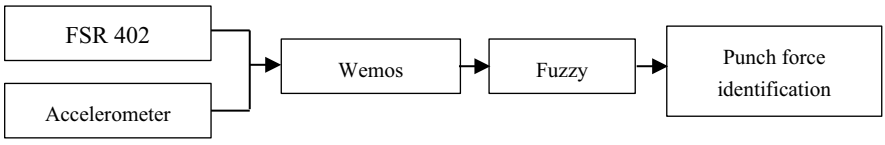


Fig. 1. System design.

The input of this system is the force variable generated from the FSR 402 sensor and the acceleration generated from the accelerometer sensor. The FSR sensor is placed on the glove so that when the athlete hits any part, the root force is measured. While the accelerometer sensor is placed on the top of the punching bag. When the punching

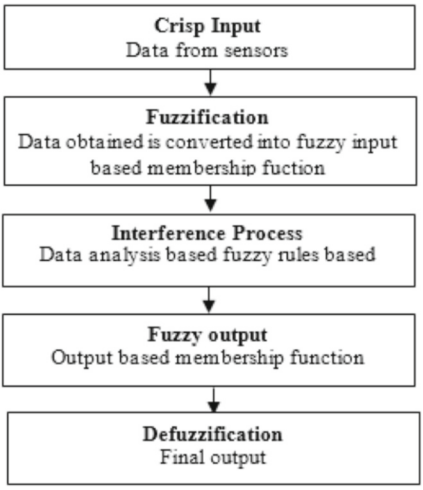


Fig. 2. Fuzzy logic design.

bag is hit, it will produce a punching bag movement, this movement will be detected so that it provides information about the acceleration of the direction of the punching bag movement.

FSR 402 and accelerometer are connected to wemos, by getting signals from these sensors, the punching force data will be obtained. Furthermore, the signals from the FSR and accelerometer sensors are processed using Fuzzy, so that the punch level is known. The fuzzy design is depicted in Fig. 2, which begins with analyzing input data and continues with the fuzzification process by converting the original input data into fuzzy form based on the membership function. The next process is the interference process and the defuzzification process. The use of fuzzy is because fuzzy has been successfully used in the assessment and evaluation process of the learning process [24–28].

Data collection was carried out 80 times to see the characteristics of the strokes. This data is categorized into 3 categories: low, medium, and high. This is a consideration in determining the membership function and also the rule-based applied in this system. Data analysis is carried out by comparing the use of both defuzzifications, and the differences between the two are discussed. Formulas of the defuzzification process implemented as follows Eq. (1) and Eq. (2):

- Centroid

$$x^* = \frac{\sum_{i=1}^{n} xi.\mu(xi)}{\sum_{i=1}^{n} \mu(xi)} \tag{1}$$

- Bisektor

$$\int_{\alpha}^{x^*} \mu_A(x)dx = \int_{x^*}^{\beta} \mu_A(x)dx \tag{2}$$

where α is defined as the minimum value of x, and β is defined as the maximum value of x, where these values satisfy the given conditions.

3 Results

The results of the design implementation are depicted in Fig. 3. It is a view of the punching bag and a picture of the electrical circuit used in this study. The glove at the end of the fist has been implanted with the FSR 402 sensor, which aims to measure the magnitude of the force that hits the punching bag. While the accelerometer sensor is placed on the top of the punching bag.

Fig. 3. Results of the implementation of the hardware design

The first and second experiments used fuzzy mamdani as depicted in Fig. 4. The input variables in this system are FSR force (force received by the sensor) and acceleration received by the accelerometer sensor. The output of this system is the force level of the punch imposed on the punching bag.

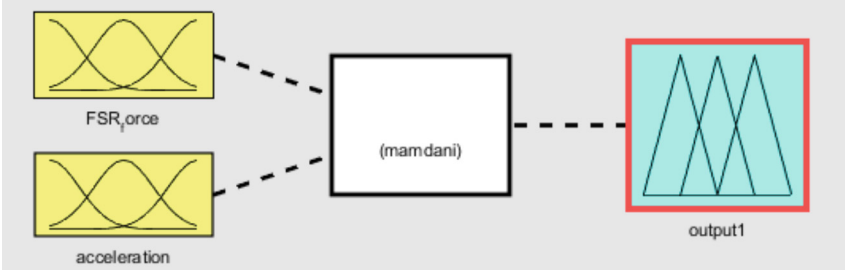


Fig. 4. Input output system design

The membership function used in this study uses a triangle as in Fig. 5. The FSR force value is between 0 and 150 Newton, and is divided into 3 (low, medium, and high). The second input variable is acceleration where the value is 0 to 20 km/s^2, divided into

3 members also (low, medium, and high). The output variable is the level of the blow which has a level of 0 to 5, just like the input variable membership function also uses a triangle with 3 members (low, medium, and high).

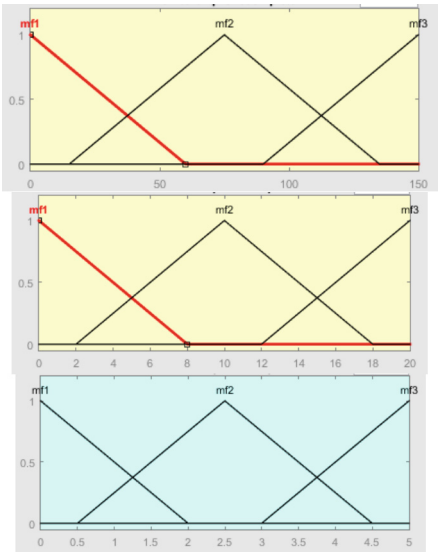


Fig. 5. Membership function of input and output variables using triangle

The rule-based system applied is depicted in Fig. 6, where in this rule the respective values for both input variables are visible. Based on this rule-based, the output can be determined. The determination of this rule is in accordance with the data analysis during the data collection process.

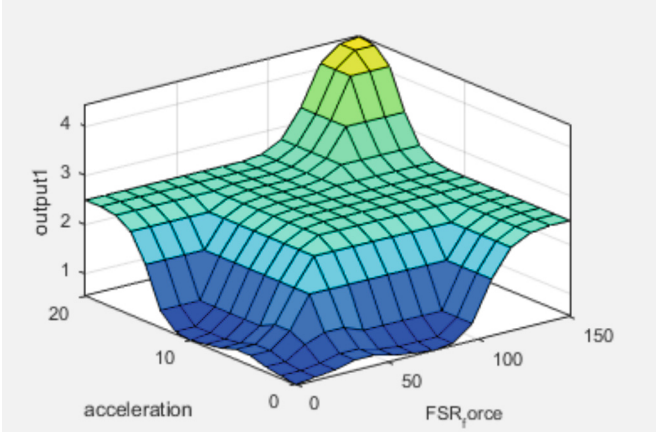


Fig. 6. Rule-based system using fuzzy mamdani

The testing process is carried out twice but using the same input variable values. This is done so that the output of both experiments can be analyzed if there is a difference in output between the two. The input values of the two variables are depicted in Fig. 7.

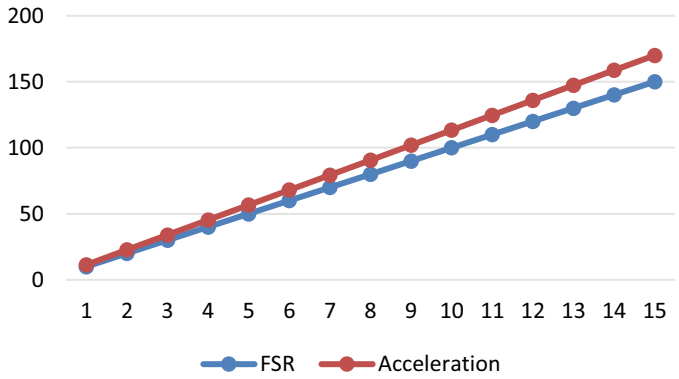


Fig. 7. Input values for the testing process in the first and second experiments.

Detailed images of the results of experiments 1 and 2 are explained in Table 1. If viewed briefly, the results of experiments 1 and 2 are almost the same, except that there are tests from 1 to 5 and also tests from 10 to 15. While for tests from 6 to 10, the output level of both experiments is the same, namely 2.5. This proves that the selection of defuzzification affects the output value.

The output results when described in detail are depicted in Fig. 8. The blue line represents the results of the first experiment and the orange line represents the results of the second experiment.

When viewed microscopically, the difference in results is depicted in Fig. 9, the difference in the first to fourth testing is positive, and the difference in the twelfth experiment tends to be negative. The average value of the delta difference between the two experiments is −0.01175, the minimum value is −0.388, and the maximum value is 0.3645.

Table 1. Results of experiment

No	FSR (N)	Acceleration (m/s)	Exp 1 the severity of punch (0–5)	Exp 2 the severity of punch (0–5)	Delta Exp 1 and 2
1	10	1.3333	0.6679	0.6	0.0679
2	20	2.6663	1.1645	0.8	0.3645
3	30	3.9993	1.7694	1.45	0.3194
4	40	5.3323	2.1295	2.15	−0.0205
5	50	6.6653	2.3573	2.4	−0.0427
6	60	7.9983	2.4992	2.5	−0.0008
7	70	9.3313	2.5	2.5	0
8	80	10.6643	2.5	2.5	0
9	90	11.9973	2.5	2.5	0
10	100	13.3303	2.6379	2.6	0.0379
11	110	14.6633	2.8601	2.85	0.0101
12	120	15.9963	3.1953	3.55	−0.3547
13	130	17.3293	3.812	4.2	−0.388
14	140	18.6623	4.3306	4.4	−0.0694
15	150	19.9953	4.35	4.45	−0.1
Defuzzification			Centroid	Bisector	
Membership function			Triangle	Triangle	
Fuzzy			Mamdani	Mamdani	
Average value					−0.01175
Minimum value					−0.388
Maximum value					0.3645

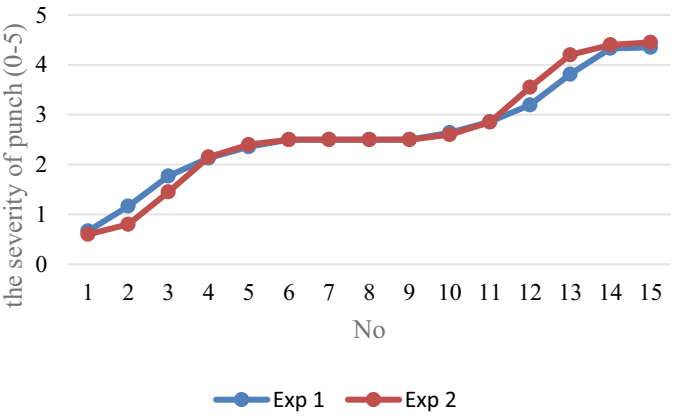


Fig. 8. Output using defuzzification centroid and bisector

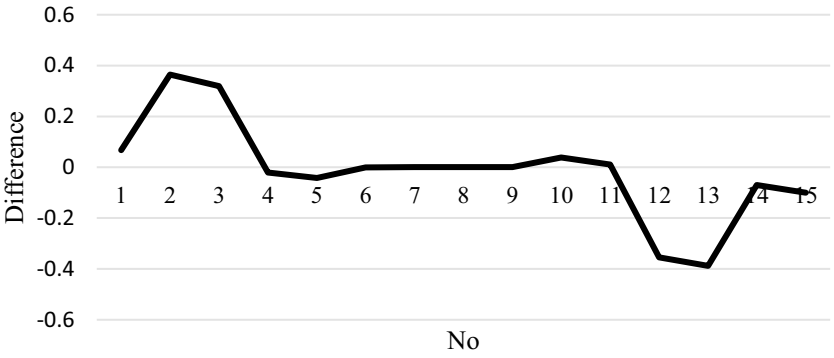


Fig. 9. Output difference using centroid and bisector defuzzification

4 Discussion

The reason why these 2 sensors are used is that in addition to having been successfully used in previous studies, they also ensures that if a punch is made with a certain force, there is a reaction, namely the movement of the punching bag that has been hit with a certain acceleration and direction. The experiment was carried out 2 times, where the first experiment used mamdani with defuzzification using centroid and the second experiment used bisector defuzzification. Centroids of fuzzy sets [29–32] and bisectors [33–35] with consideration of this method have been successfully used in previous studies. The first and second experiments used fuzzy mamdani where this method has also been implemented in several cases [36–38].

Based on Fig. 7 data analysis during the data collection process, these two variables have a tendency to be directly proportional, namely if the FRS sensor value increases, there is a tendency for the acceleration value to also increase. Fig. 8 shows there is a difference between the two where this difference is found in the first to fifth tests and

the eleventh to fifteenth tests. While the results of the sixth and tenth tests, the blue and orange lines overlap, this shows that this test produces the same results.

Until now, research on punching bag technology innovation has rarely been carried out, even though the technology is needed by society. Previous research on punching bags is an innovative punching bag system in which a micro-energy generator is embedded to produce energy, so it can be an alternative renewable energy source. This research is more concentrated on how to monitor and know the progress of hitting training done by athletes. So it is known how many levels of style have been done. If athletes know this condition, athletes can find ways to improve their performance. Research [16, 17] Has been able to analyze punching movements in martial sports but does not discuss the magnitude of the force that has been applied. With this research, it can complement the analysis of movements and punching force can also be analyzed.

The weakness of this system is that it depends on the specifications of the sensor used, in this case, it is limited to a maximum of 150 N. The advantage of this system is that the user can hit any part of the punching bag, whereas previous research was limited to only one point. With this advantage, the user can be more flexible in hitting the punching bag. The implication of this research is that users can analyze the strokes that have been made in more detail so that actions can be taken according to the analysis results so that performance can be improved.

5 Conclusion

The purpose of this study is to propose a smart punching bag to identify the force level of the punches that have been performed using fuzzy. Based on the experiments that have been conducted, the proposed approach punching bag for punch force level identification in self-defense training using fuzzy logic can be implemented, although using different defuzzification parameters produces different output results. However, this difference is not too large, namely the average value of the delta difference between the two experiments is −0.01175, the minimum value is −0.388, and the maximum value is 0.3645. Future research is to conduct research with more than one FSR sensor variable, in addition, it can be equipped with monitoring to what extent athletes can make punches according to the specified target.

Acknowledgments. We would like to thank the Directorate General of Higher Education, Research, and Technology of the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia and Universitas Negeri Surabaya for providing funding, facilities, and infrastructure so that this research can be carried out properly.

References

1. Ciaccioni S et al (2022) The interlink among age, functional fitness, and perception of health and quality of life: a mediation analysis. Int J Environ Res Public Health 19. https://doi.org/10.3390/ijerph19116850
2. Mills K, Dudley D, Collins NJ (2019) Do the benefits of participation in sport and exercise outweigh the negatives? An academic review. Best Pract Res Clin Rheumatol 33:172–187

3. Oja P et al (2024) Health benefits of different sports: a systematic review and meta-analysis of longitudinal and intervention studies including 2.6 million adult participants. Sports Medicine Open 10:1–17
4. Biddle S, Mutrie N, Trish Gorely GF (2021) Psychology of physical activity determinants, well-being and interventions. Routledge. https://doi.org/10.4324/9781003127420
5. Dongoran MF, Fadlih AM, Riyanto P (2020) Psychological characteristics of martial sports Indonesian athletes based on categories art and fight. Enferm Clin 40:500–503
6. Ciaccioni S et al (2024) Martial arts, combat sports, and mental health in adults: a systematic review. Psychol Sport Exerc 70:1–18
7. Munir HMW, Latif A, Mehmood N, Patafi MAM, Ahmed N, M. (2023) Effect of punching bag exercises on body mass index (BMI). J Health Rehabilitation Res 3:965–969
8. Anifah L, Syariffuddien Zuhrie M, Muhammad, Haryanto (2024) A novel pencak silat punch pattern recognition approach using supervised learning. Ain Shams Eng J 15:102857. https://doi.org/10.1016/j.asej.2024.102857
9. Labintsev A, Khasanshin I, Balashov D, Bocharov M, Bublikov K (2021) Recognition punches in karate using acceleration sensors and convolution neural networks. IEEE Access 9:138106–138119. https://doi.org/10.1109/ACCESS.2021.3118038
10. Omcirk D, Vetrovsky T, Padecky J, Vanbelle S, Malecek J, Tufano JJ (2021) Punch trackers: correct recognition depends on punch type and training experience. Sensors 21:1–10. https://doi.org/10.3390/s21092968
11. Kasiri S, Fookes C, Sridharan S, Morgan S (2017) Fine-grained action recognition of boxing punches from depth imagery. Comput Vis Image Underst 159:143–153. https://doi.org/10.1016/j.cviu.2017.04.007
12. Khasanshin I, Osipov A (2021) Using an artificial neural network to develop an optimal model of straight punch in boxing and training in punch techniques based on this model and real-time feedback. PLoS One 16:1–17. https://doi.org/10.1371/journal.pone.0259457
13. Chaithrai D, Babu D, Malligei B, Roopashreei S (2020) Smart punching bag system with micro energy generator. Int J Adv Sci Technol 29:8518–8525
14. Lapkova D, Adamek M (2015) Using strain gauge for measuring of direct punch force. In: XXI IMEKO World Congress "Measurement in Research and Industry", Prague, Czech Republic
15. Lapkova D, Pluhacek M, Adamek M (2014) Numerical analysis of direct punch with a view to velocity and level of training. In: Proceedings of the international conference on numerical analysis and applied mathematics 2014 (ICNAAM-2014)
16. Anifah L, Syariffuddien Zuhrie M, Muhammad, Haryanto (2024) A novel pencak silat punch pattern recognition approach using supervised learning. Ain Shams Eng J 15:102857. https://doi.org/10.1016/j.asej.2024.102857
17. Khasanshin I, Osipov A (2021) Using an artificial neural network to develop an optimal model of straight punch in boxing and training in punch techniques based on this model and real-time feedback. PLoS One 16:1–17. https://doi.org/10.1371/journal.pone.0259457
18. Velásquez EIG, Gómez V, Paredes-Madrid L, Colorado HA (2019) Error compensation in force sensing resistors. Sensing Bio-Sensing Res 26
19. Biju SM, Sheikh HZ, Malek MF, Oroumchian F, Bell A (2021) Design of grip strength measuring system using FSR and flex sensors using SVM algorithm. IAES Int J Artif Intell (IJ-AI) 10(676–686)
20. Biddle S, Mutrie N, Trish Gorely GF (2021) Psychology of physical activity determinants, well-being and interventions. Routledge. https://doi.org/10.4324/9781003127420
21. Oyubu AO, Mbachu CB, Nwabueze CA, Iloh JP (2022) A force sensitive resistor based wireless sensor network for pipeline monitoring and oil spillage control in Nigeria. Eur J Eng Technol Res 7

22. Erdaş Ç, Atasoy I, Açıcı K, Oğul H (2016) Integrating features for accelerometer-based activity recognition. In: The 3rd international symposium on emerging information, communication and networks (EICN 2016)
23. Rahaman MM, Fang W, Fawzi AL, Wan Y, Kesari H (2020) An accelerometer-only algorithm for determining the acceleration field of a rigid body, with application in studying the mechanics of mild traumatic brain injury. J Mech Phys Solids 143
24. Imam I (2017) Obtaining ABET student outcome satisfaction from course learning outcome data using fuzzy logic. EURASIA J Math Sci Technol Educ 13:3069–3081
25. Ghomshei: Fuzzy logic in engineering education and evaluation of graduate attributes. In: Proceedings of 2018 Canadian Engineering Education Association (CEEA-ACEG18) conference (2018)
26. Voskoglou (2013) Fuzzy logic as a tool for assessing students' knowledge and skills. Educ Sci (Basel) 3:208–221
27. Thomas (2019) A novel fuzzy based assessment method for learning outcomes. Int J Recent Technol Eng (IJRTE) 8:1665–1669
28. Venkatesan: Evaluating learning outcomes in PBL using fuzzy logic techniques. In: Proceedings of the 2008 AaeE conference (2008)
29. Aisbett J, Rickard JT (2013) Centroids of fuzzy sets when membership functions have spikes. In: 2013 joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), Edmonton, AB, Canada, pp 97–101. https://doi.org/10.1109/IFSA-NAFIPS.2013.6608382
30. Kamis NH, Abdullah K, Noh NAM, Bakar NAA, Saad RM, Saat SZ (2012) Centroid method in fuzzy LinPrera: a case study of UiTM, Shah Alam. In: 2012 IEEE colloquium on humanities, science and engineering (CHUSER), Kota Kinabalu, Malaysia, pp 544–548
31. Aisbett J, Rickard JT (2014) Centroids of type-1 and type-2 fuzzy sets when membership functions have spikes. IEEE Trans Fuzzy Syst 22:685–692. https://doi.org/10.1109/TFUZZ.2014.2306973
32. Figueroa–García JC, Orjuela–Cañon AD, Neruda R (2019) On the boundaries of the centroid of a class of fuzzy numbers. In: 2019 IEEE Latin American Conference on Computational Intelligence (LA-CCI), pp 1–5. https://doi.org/10.1109/LA-CCI47412.2019.9037055.
33. Mogharreban NDLF (2006) Comparison of defuzzification techniques for analysis of non-interval data. In: NAFIPS 2006–2006 annual meeting of the north american fuzzy information processing society, pp 257–260. https://doi.org/10.1109/NAFIPS.2006.365418
34. Lukács J (2019) Comparison of defuzzification methods for cabin noise prediction of passenger cars. In: 2019 IEEE 17th international symposium on intelligent systems and informatics (SISY), Subotica, Serbia, pp 000115–000120. https://doi.org/10.1109/SISY47553.2019.9111606
35. Pasi Luukka JS (2024) Possibilistic mean based defuzzification for fuzzy expert systems and fuzzy control—LSD for general fuzzy sets. J Comput Appl Math 441. https://doi.org/10.1016/j.cam.2023.115663
36. Lilly JH (2010) Mamdani fuzzy systems. Fuzzy control and identification, pp 27–45. https://doi.org/10.1002/9780470874240.ch3
37. Sakti I (2014) Methodology of fuzzy logic with mamdani fuzzy models applied to the microcontroller. In: 2014 the 1st international conference on information technology, computer, and electrical engineering, Semarang, Indonesia, pp 93–98. https://doi.org/10.1109/ICITACEE.2014.7065721
38. Nanda AR, Rathanda AK (2017) Mamdani fuzzy inference based hierarchical cost effective routing (MFIHR) in WSNs. In: 2017 IEEE 7th international advance computing conference (IACC), pp 397–401. https://doi.org/10.1109/IACC.2017.0089

Optimation Environment Estimation for Maximum Fertility Level of Zea mays ssp. Plants in Automatic Fodder Using Fuzzy Logic

Haryanto[1(✉)], Dwi Kuswanto[2], Dian Neipa Purnamasari[1], and Lilik Anifah[3]

[1] Department of Electrical Engineering, Faculty of Engineering, Universitas Trunojoyo Madura, Jawa Timur, Indonesia
haryanto@trunojoyo.ac.id

[2] Department of Informatic Engineering, Faculty of Engineering, Universitas Trunojoyo Madura, Jawa Timur, Indonesia

[3] Department of Electrical Engineering, Faculty of Engineering, Universitas Negeri Surabaya, Jawa Timur, Indonesia

Abstract. Animal food is very important to achieve the availability of animal protein for the population. If the availability of livestock food is limited, there will be an impact on the availability of food for humans. Some areas experiencing difficulties in animal feed in the dry season can be overcome with fodder using Zea mays ssp. seeds. The purpose of this research is to estimate the environmental optimization conditions needed for Zea mays ssp. to grow optimally. The research stages are hardware development, data collection, fuzzy design process, and testing process. The sensor used as fuzzy input is ultrasonic, this sensor is used to monitor the height growth of Zea mays ssp. The signals produced by these sensors become input for fuzzy, which is then used to estimate the optimum environmental conditions of the fodder so that the growth of corn plants on the fodder can be improved. Optimization environment estimation for the maximum fertility level of Zea mays ssp. plants in automatic fodder using fuzzy logic has been applied, based on the test results the optimum result is when a water level of 20 mm and light intensity of 0 to 500 lx are used. Another optimum condition is when given the condition of 30 mm water level with light intensity of 0, 200, 300, and 500 lx. These two conditions obtained a fertility level of 0.8628 when using centroid de-fuzzification and 0.88 when using bisector defuzzification, where the average of both is 0.8714. The contribution of this research is expected to help farmers in determining optimal environmental variables so that Zea mays ssp. plants can grow well.

Keywords: Optimation · Environment · Estimation · Zea mays ssp. · Fuzzy

1 Introduction

Several regions in Indonesia experience difficulties in animal feed in the dry season. During the dry season, farmers in the Belu Regency of East Nusa Tenggara are generally faced with the problem of difficulty in fulfilling animal feed [1]. The same thing also

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 291–304, 2026.
https://doi.org/10.1007/978-981-95-9233-3_22

happened in Gorontalo, extreme hot weather due to the phenomenon of El Nino sea level temperature warming in Gorontalo Province is also quite felt by cattle herders [2], also occurred in Bali [3], Semarang [4], and may occur in other regions in Indonesia. If this happens for a long time, it is feared that it will disadvantage farmers.

Some solutions offered by previous research are bio-fermentation technology or silage technology [5, 6]. Silage can be made from dry rice straw and processed into fermented rice straw, which is turned into complete feed by adding other ingredients so that its quality increases, from three to four percent protein content to seven to eight, and is resistant throughout the season. Ingredients used in fermented complete feed technology, such as rice straw and several types of concentrate feed ingredients (optional and situational) include rice bran, onggok, cassava, copra meal, peanut shells, bread, ground yellow corn, and microbial starters. Another alternative is to provide food in the form of microgreens produced from fodder. The problem is that until now there is no automatic fodder so as to improve the productivity of farmers. This research aims to design an automatic smart fodder system as an animal feed production machine as a solution to the difficulty of animal feed. It is expected that this technology will be useful for farmers and can increase their productivity.

Feed shortages during the dry season are due to a lack of rain, resulting in dry grass. The problem-solving approach is to find an alternative solution by producing microgreen corn using fodder. This is supported by data that for two weeks, the weight gain of 1 kg of seeds can produce about 10 kg of corn fodder. The results of the analysis of variance testing the nutrient content of corn fodder in the form of the percentage of dry matter (BK), organic matter (BO), crude protein (PK), crude fiber (SK), and crude fat (LK) showed that the harvesting age of 8 HST, 10 HST, 12 HST and 14 HST gave a significant effect on the nutrient content of corn fodder on the percentage of dry matter (BK), organic matter (BO), crude protein (PK) and crude fat (LK) with ($P < 0.05$) [7]. This has been shown to have a good impact on livestock [8]. The next problem-solving approach is to design an Automatic Smart Fodder System as an animal feed production machine so that it can be efficient in terms of labor and cost. The first approach is to use a water drop sensor, where this sensor functions to ensure that all parts of the microgreen have been watered [9]. It will also keep the system water-efficient. This system is also equipped with a timer where the system will do watering periodically, thus saving the farmer's labor. Also equipped with ultrasonic that can identify the height of the microgreen. It is also provided with an LDR to anticipate the intensity of light entering the system. The important thing is also equipped with a water sensor on the water storage, and there will be a notification when the water runs out.

Fuzzy has been used in agriculture, including the use of fuzzy has been used for irrigation management [10]. Implementation in tropical agriculture has also been used and proven to make agriculture more efficient [11]. The use of fuzzy combined with wireless sensors in agriculture has become more efficient [12]. Fuzzy is also used for indoor irrigation and has been proven to help in the automatic control of irrigation systems [13]. Fuzzy has also been successfully applied to smart farming [14]. Based on this, this system uses fuzzy in finding optimum environmental conditions in the automatic fodder system.

Previous research on farming has been investigated by several researchers before. Smart aquaponic system based on Internet of Things (IoT) [15] has been studied but this system is not specific about fodder. Research on fodder is designed to build an IoT-based monitoring system for fodder [16]. A 1-level fodder system has been built [17], but in this system, there is only 1 level so that the results are still less efficient. Specific research on fodder is still limited, but considering the availability of animal feed is needed for farmers, especially during the dry season, this research is important and urgent to do. Based on previous research, fuzzy has been successfully applied in the agricultural sector.

This research aims to design optimization environment control for maximum fertility level of Zea mays ssp. plants in automatic fodder using fuzzy logic so that farmers help farmers in conditioning plants so that maximum results are obtained. The hope of this research is that the proposed system can be useful for farmers and become a solution for providing feed during the dry season. The originality of this research is that this system as an animal feed production machine is equipped with a water level sensor to monitor sprinkler water, a water drop to ensure that all parts are exposed to water, ultrasonic, and also an LDR sensor (light sensor). In this system, the environment is set so that Zea mays ssp. plants can grow well.

2 Methods

The research stages are hardware development, data collection, fuzzy design process, and testing process. The hardware referred to in this research is the sensor, NodeMCU and all fodder devices so that Zea Mays can grow well. The sensor used as fuzzy input is ultrasonic (HC-SR04), this sensor is used to monitor the height growth of Zea mays ssp. In addition, an LDR sensor (NORP12 RS) is used to monitor fodder lighting, and a water level sensor (K-0135) to monitor water levels. The signals generated by these sensors become input for fuzzy, to then estimate the optimum environmental conditions of the fodder so that the growth of corn plants on fodder can be improved.

The ultrasonic sensor (HCSR04) is a distance sensor that works by sending waves to an object and receiving the signal back [18–20]. This sensor has been widely used in various fields, previous research used it for Unmanned Aerial Vehicle (UAV) [21], material classification [22], social distance vest [23], etc. Light Dependent Resistor (LDR) is a sensor that can measure light intensity [24, 25], this sensor has been widely used in previous research [26].

Table 1. Ideal growth height of corn [27]

Age (day)	Height (cm)	Standard deviation (+/− cm)
8	11.44	3.26
10	15.97	6.3
12	19.12	4.92
14	26.92	1.37

The growth of corn plants from 8 to 14 days old is described in Table 1. Monitoring in this study was carried out when the corn plants were 8, 10, 12, and 14 days old. The fuzzy design implemented uses Mamdani fuzzy with a defuzzification process using a centroid and bisector. The defuzzification function is where the fuzzy results are changed into one value [27, 28]. The formulation of centroid and bisector is explained in Eq. (1) for centroid and Eq. (2) for bisector [29].

$$x^* = \frac{\sum_{i=1}^{n} xi.\mu(xi)}{\sum_{i=1}^{n} \mu(xi)} \tag{1}$$

$$\int_{\alpha}^{x^*} \mu_A(x)dx = \int_{x^*}^{\beta} \mu_A(x)dx \tag{2}$$

where α is defined as the minimum value of x, and β is defined as the maximum value of x, where these values satisfy the given conditions. Experiments were conducted by testing the results of fuzzy, then continued with the search for optimum environmental conditions according to the needs of corn plants. The fuzzy formation process becomes 4 stages, namely the stage when the plant is 8 days old, 10 days old, 12 days old, and 14 days old. The type used is mamdani fuzzy with a triangular membership function, and defuzzification using centroid and bisector.

This study uses the fuzzy method so that data collection is carried out as a basis for determining the membership function of each variable. Data collection is carried out directly 4 times, where each data is taken 25 times so that the data obtained is 100 data. Based on this data, the membership function and also the rules-based system are determined. The analysis is carried out by comparing the output produced by fuzzy and the actual data. In this way, the difference between the results of this system and the actual conditions is obtained.

3 Results

One of the reasons why fuzzy is used is because it has been successfully applied in previous research [30, 31], including implementation in the agricultural sector [32–34]. The proposed fuzzy picture is depicted in Fig. 1. The input of the system is the height of the plant generated from the ultrasonic sensor, water level, and light intensity where this information is obtained from the LDR sensor. The input of these three variables will produce a value of the fertility level of the plant.

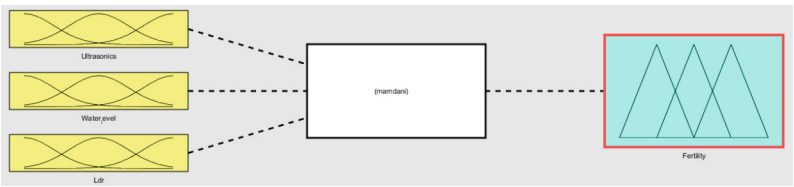


Fig. 1. Design of fuzzy used in experiment

The membership functions of the input and output variables are depicted in Fig. 2. The four experiments for plant conditions aged 8 days, 10 days, 12 days, and 14 days using a triangular membership function. The number of sets is 3 (low, medium, and high) for all input and output variables. The difference is that each input and output variable has its own minimum and maximum values.

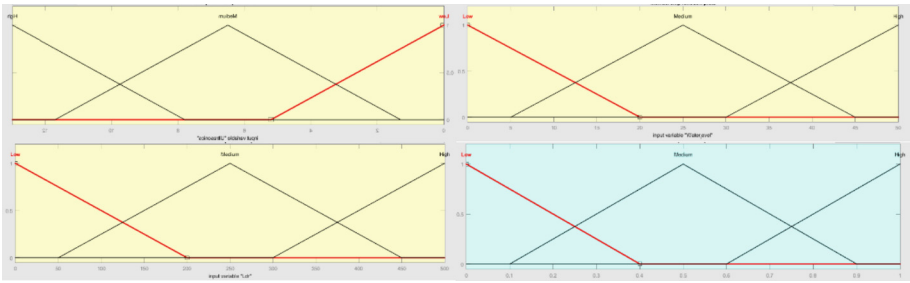


Fig. 2. Membership function of fuzzy system

Rules-based surface visualization of the system applied in this study is depicted in Fig. 3. Rules-based is applied to all experiments, this is conducted so that all experiments are ensured to have the same rules. With this treatment, the optimization carried out obtained more fair results for all values applied. There are 4 rules-based applied in the research, namely for plant conditions aged 8 days, 10 days, 12 days, and 14 days. The description of one of the rules-based for plant conditions aged 8 days is as follows:

If (Ultrasonics is Low) and (Water_level is Low) and (Ldr is Low) then (Fertility is Low)
If (Ultrasonics is Medium) and (Water_level is Low) and (Ldr is Low) then (Fertility is Low)
If (Ultrasonics is High) and (Water_level is Low) and (Ldr is Low) then (Fertility is Medium)
If (Ultrasonics is Low) and (Water_level is Low) and (Ldr is Low) then (Fertility is Low)
If (Ultrasonics is Low) and (Water_level is Medium) and (Ldr is Low) then (Fertility is Low)
If (Ultrasonics is Medium) and (Water_level is Medium) and (Ldr is Low) then (Fertility is Medi-um)
If (Ultrasonics is High) and (Water_level is Medium) and (Ldr is Low) then (Fertility is High)
If (Ultrasonics is Low) and (Water_level is High) and (Ldr is Low) then (Fertility is Medium)
If (Ultrasonics is Medium) and (Water_level is High) and (Ldr is Low) then (Fertility is Medium)
If (Ultrasonics is High) and (Water_level is High) and (Ldr is Low) then (Fertility is Medium)
If (Ultrasonics is Low) and (Water_level is Low) and (Ldr is Medium) then (Fertility is High)
If (Ultrasonics is Medium) and (Water_level is Medium) and (Ldr is Medium) then (Fertility is Medium)

If (Ultrasonics is High) and (Water_level is Medium) and (Ldr is Medium) then (Fertility is High)
If (Ultrasonics is Low) and (Water_level is High) and (Ldr is Medium) then (Fertility is Low)
If (Ultrasonics is Medium) and (Water_level is High) and (Ldr is Medium) then (Fertility is Medi-um)
If (Ultrasonics is High) and (Water_level is High) and (Ldr is Medium) then (Fertility is Medium)
If (Ultrasonics is Low) and (Water_level is Low) and (Ldr is High) then (Fertility is Low)
If (Ultrasonics is Medium) and (Water_level is Low) and (Ldr is High) then (Fertility is Medium)
If (Ultrasonics is High) and (Water_level is Low) and (Ldr is High) then (Fertility is Medium)
If (Ultrasonics is Low) and (Water_level is Medium) and (Ldr is High) then (Fertility is Low)
If (Ultrasonics is Medium) and (Water_level is Medium) and (Ldr is High) then (Fertility is Medi-um)
If (Ultrasonics is High) and (Water_level is Medium) and (Ldr is High) then (Fertility is High)
If (Ultrasonics is Low) and (Water_level is High) and (Ldr is High) then (Fertility is Low)
If (Ultrasonics is Medium) and (Water_level is High) and (Ldr is High) then (Fertility is Medium)
If (Ultrasonics is High) and (Water_level is High) and (Ldr is High) then (Fertility is Medium)

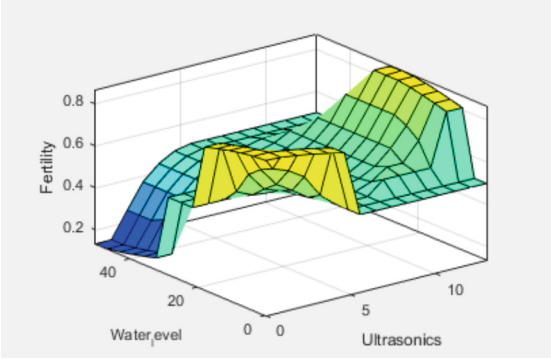


Fig. 3. Rules-based system surface visualization

Tests for each experiment were conducted 50 times, examples of the results obtained for the four experiments are described in Table 2, Table 3, Table 4, and Table 5. This data was obtained from data collection from the growth process in automatic fodder. Table 2 shows that the defuzzification results using centroid and bisector show a high

difference of −0.08 to 0.05. Plant fertility levels are expressed with values of 0 to 1, where 0 indicates infertile while 1 indicates very fertile.

Table 2. Examples of some Fuzzy test data on the fodder system Zea mays ssp. at 8 days

No	Ultrasonics (cm)	Water level (mm)	LDR (lux)	Defuzzification Centroid	Defuzzification Bisector	Delta centroid and bisector
1	9	10	100	0.41	0.36	0.05
2	10	20	500	0.6	0.62	−0.02
3	11	30	400	0.7	0.78	−0.08
4	12	40	300	0.54	0.54	0
5	13	50	200	0.5	0.5	0

Table 3 shows the test results for 10-day-old plants. The difference using centroid and bisector defuzzification is −0.1 to 0.18, which means that using these two types of defuzzification shows almost the same results.

Table 3. Examples of some Fuzzy test data on the fodder system Zea mays ssp. at 10 days

No	Ultrasonics (cm)	Water level (mm)	LDR (lux)	Defuzzification Centroid	Defuzzification Bisector	Delta centroid and bisector
1	9	10	100	0.41	0.29	0.12
2	10	20	500	0.6	0.5	0.1
3	11	30	400	0.7	0.52	0.18
4	12	40	300	0.54	0.54	0
5	13	50	200	0.5	0.5	0

Table 4 shows the difference using centroid and bisector defuzzification is higher than the 10-day-old system. The difference is in the range of 0 to 0.3428, the highest difference occurs when using a water level of 30 mm and a light intensity of 300 lx with a plant height of 8 cm. However, the other data results did not show a big difference.

The difference in the use of both fuzzifications is in the range of −0.0869 to 0.0086, where the highest difference is the lowest in this system. The best conditioning is when a water level of 20 mm and light intensity of 200 lx is applied, with a plant fertility rate of 0.8628 when using centroid and 0.88 when using bisector defuzzification.

After initial testing, based on the above system, the optimum environmental conditions were determined for maximum plant growth. The results of this test are described in Fig. 4, Fig. 5, Fig. 6, and Fig. 7. In those figures, the y-axis shows the fertility level of the plants, while the x-axis is the number of experiments that have been carried out.

Table 4. Examples of some Fuzzy test data on the fodder system Zea mays ssp. at 12 days

No	Ultrasonics (cm)	Water level (mm)	LDR (lux)	Defuzzification Centroid	Defuzzification Bisector	Delta centroid and bisector
1	6	10	100	0.5	0.29	0.21
2	7	20	200	0.7292	0.51	0.2192
3	8	30	300	0.8628	0.52	0.3428
4	9	40	400	0.5486	0.54	0.0086
5	10	50	500	0.5	0.5	0

Table 5. Examples of some Fuzzy test data on the fodder system Zea mays ssp. at 14 days

No	Ultrasonics (cm)	Water level (mm)	LDR (lux)	Defuzzification Centroid	Defuzzification Bisector	Delta centroid and bisector
1	23	30	300	0.7331	0.82	−0.0869
2	24	40	400	0.5486	0.54	0.0086
3	25	50	500	0.5	0.5	0
4	26	10	100	0.5486	0.54	0.0086
5	27	20	200	0.8628	0.88	−0.0172

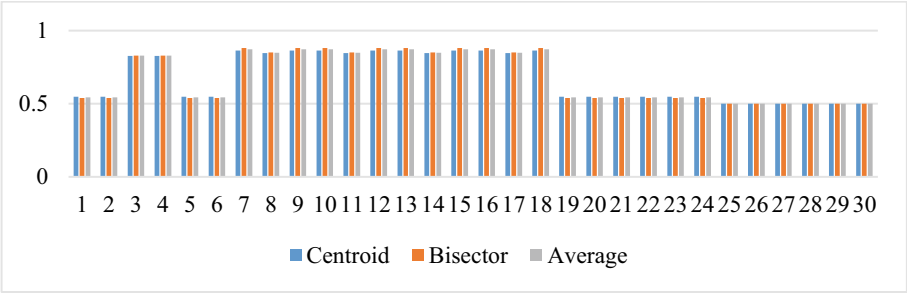


Fig. 4. Visualization of fertility levels for Zea mays ssp. at 8 days

The results of the optimum values for water level and light intensity for maximum fertility level of Zea mays ssp. plants in automatic fodder using fuzzy logic are described in Table 6 for 8-day-old plants, Table 7 for 10-day-old plants, Table 8 for 12-day-old plants, and Table 9 for 14-day-old plants.

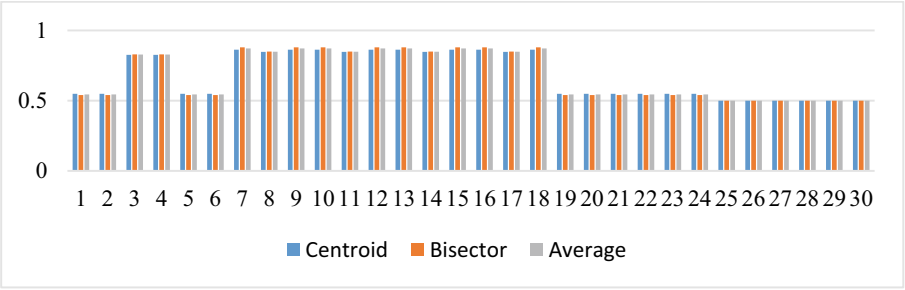


Fig. 5. Visualization of fertility levels for Zea mays ssp. at 10 days

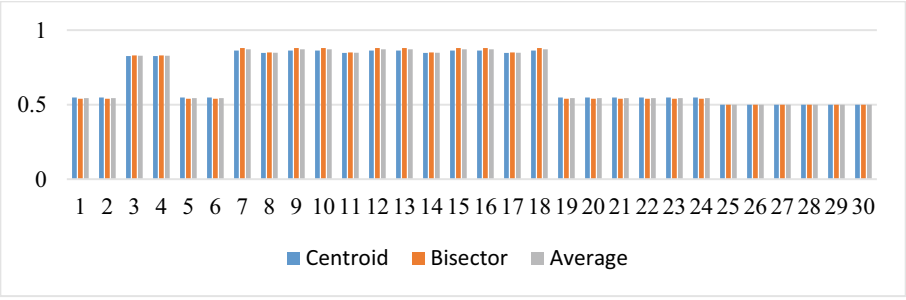


Fig. 6. Visualization of fertility levels for Zea mays ssp. at 12 days

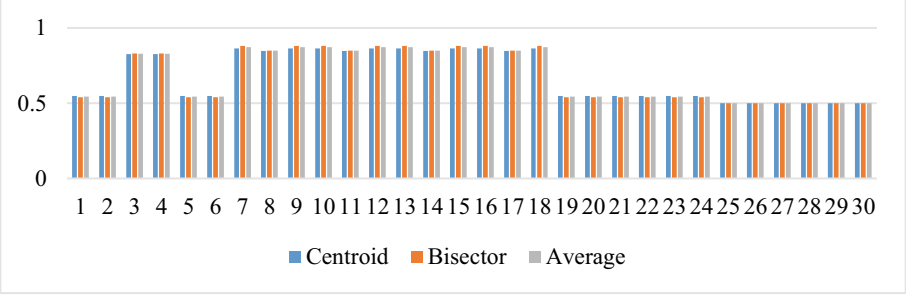


Fig. 7. Visualization of fertility levels for Zea mays ssp. at 14 days old

The optimum condition for the entire system is obtained with a pair of conditions, where the optimum condition is when a water level of 20 mm and light intensity of 0 to 500 lx are used. Another optimum condition is when given the condition of 30 mm water level with light intensity of 0, 200, 300, and 500 lx.

Based on the results of this experiment, fodder should be conditioned at water levels of 20 mm and 30 mm. The experiments show that this water level produces a plant fertility level above 0.6. Meanwhile, if the system is applied with a water level of 0 to 10, the maximum plant fertility level is 0.6, which means that this condition should be

Table 6. Results of optimum values for water level and light intensity for maximum fertility level of Zea mays ssp. plants for 8 days of age

No	Ultrasonics (cm)	Water level (mm)	LDR (lux)	Defuzzification Centroid	Defuzzification Bisector	Delta centroid and bisector
1	13	20	0	0.8628	0.88	0.8714
2	13	20	200	0.8628	0.88	0.8714
3	13	20	300	0.8628	0.88	0.8714
4	13	20	500	0.8628	0.88	0.8714
5	13	30	0	0.8628	0.88	0.8714
6	13	30	200	0.8628	0.88	0.8714
7	13	30	300	0.8628	0.88	0.8714
8	13	30	500	0.8628	0.88	0.8714

Table 7. Results of optimum values for water level and light intensity for maximum fertility level of Zea mays ssp. plants for 10 days of age

No	Ultrasonics (cm)	Water level (mm)	LDR (lux)	Defuzzification Centroid	Defuzzification Bisector	Delta centroid and bisector
1	17	20	0	0.8628	0.88	0.8714
2	17	20	200	0.8628	0.88	0.8714
3	17	20	300	0.8628	0.88	0.8714
4	17	20	500	0.8628	0.88	0.8714
5	17	30	0	0.8628	0.88	0.8714
6	17	30	200	0.8628	0.88	0.8714
7	17	30	300	0.8628	0.88	0.8714
8	17	30	500	0.8628	0.88	0.8714

avoided so that farmers do not experience losses. Water levels that are too low will cause plants to lack nutrients, resulting in disrupted growth.

The condition of using a water level above 30 mm also has a high risk, proven by the maximum plant fertility level of 0.6. Conditions in the field if the water level is too high, the plant will cause the plant to rot, characterized by brownish seeds.

Table 8. Results of optimum values for water level and light intensity for maximum fertility level of Zea mays ssp. plants for 12 days old

No	Ultrasonics (cm)	Water level (mm)	LDR (lux)	Defuzzification Centroid	Defuzzification Bisector	Delta centroid and bisector
1	20	20	0	0.8628	0.88	0.8714
2	20	20	200	0.8628	0.88	0.8714
3	20	20	300	0.8628	0.88	0.8714
4	20	20	500	0.8628	0.88	0.8714
5	20	30	0	0.8628	0.88	0.8714
6	20	30	200	0.8628	0.88	0.8714
7	20	30	300	0.8628	0.88	0.8714
8	20	30	500	0.8628	0.88	0.8714

Table 9. Results of optimum values for water level and light intensity for maximum fertility level of Zea mays ssp. plants for 14 days

No	Ultrasonics (cm)	Water level (mm)	LDR (lux)	Defuzzification Centroid	Defuzzification Bisector	Delta centroid and bisector
1	27	20	0	0.8628	0.88	0.8714
2	27	20	200	0.8628	0.88	0.8714
3	27	20	300	0.8628	0.88	0.8714
4	27	20	500	0.8628	0.88	0.8714
5	27	30	0	0.8628	0.88	0.8714
6	27	30	200	0.8628	0.88	0.8714
7	27	30	300	0.8628	0.88	0.8714
8	27	30	500	0.8628	0.88	0.8714

4 Discussion

The relationship between ultrasonics, water level, and LDR with fertility level is if the lighting (which can be measured through the LDR sensor) and the level of water in the Zea mays growing area (which is measured by the water level sensor) are in accordance with the needs of the corn plant, then the plant will grow well where its height can be measured through the ultrasonic sensor). While the correlation to fertility fuzzy membership function is a signal received from the three sensors above, its rules-based system is made according to the needs of environmental variables and the fertility

level of the plant where the value is between 0 and 1 (0 indicates infertile, while 1 indicates very fertile).

This research is useful for farmers, if farmers use the manual fodder system, it is recommended to use sufficient water, which is marked in this study using a water level between 20 and 30 mm. Plants will grow optimally, where the condition of the plant does not lack nutrients to grow and also does not rot because of too much water. If farmers use an automatic fodder system then this system can be used as a reference to get maximum results.

Research on fodder has rarely been done before. Research has been done well, where this system is based on IoT [35], as well as in [36], but this system has not conditioned the fodder for each age of the corn plant. In fact, the optimum conditions for each age are different. This system attempts to condition the system at the ages of 8, 10, 12, and 14 days, so that if something unexpected happens (or does not match the conditions that should be) farmers can take immediate action.

Based on the results obtained, it show that both defuzzifications can be used because the application of both defuzzifications produces almost the same value of environmental variables needed by plants. However, the use of both defuzzifications obtained fertility levels with a difference of 0.8714, whereas the use of bisectors showed higher numbers. The limitation of this system is that the algorithm applied is rather complicated because every age of 8, 10, 12, and 14 days fuzzy is applied. This is natural because the ideal growth of each age is different. If fuzzy is applied to all ages, it is feared that the predicted growth will not be in accordance with the conditions of each age.

5 Conclusion

Optimization environment estimation for the maximum fertility level of Zea mays ssp. plants in automatic fodder using fuzzy logic has been applied, based on the test results the optimum result is when a water level of 20 mm and light intensity of 0 to 500 lx are used. Another optimum condition is when given the condition of 30 mm water level with light intensity of 0, 200, 300, and 500 lx. These two conditions obtained a fertility level of 0.8628 when using centroid defuzzification and 0.88 when using bisector defuzzification, where the average of both is 0.8714. This condition is supported by the fact that if the plants in the fodder are given too high water levels, the grains will rot and eventually cannot grow properly. Different conditions if the water given is too low then the plants will dry out and cannot grow due to lack of nutrients. Future work from this research is that other artificial intelligence can be applied and also the use of rules-based systems by confirming the test results on farmers.

Acknowledgments. Our thanks go to the Directorate of Research and Community Service Republic of Indonesia and Trunojoyo Madura University for supporting and facilitating our research.

References

1. Ama, K.K.: Puncak Kemarau Penyebab Bobot Sapi di NTT Turun Drastis.

2. Ibrahim A. Musim Kemarau, Peternak Sapi di Gorontalo Kesulitan Dapatkan Pakan Ternak
3. Parananda E. Musim Kemarau, Peternak Sapi Kesulitan Pakan dan air Minum
4. Muchamad Dafi Yusuf DAR. Kekeringan, Peternak di Semarang Kesulitasn Mendapatkan Rumput untuk Pakan Sapi.
5. Cakra I, Putri B, Budaarsa I, Sukmawati NMS, Wibawa A (2018) Penerapan Teknologi Silase Ransum Komplit Dalam Upaya Mengatasi kekurangan Pakan Pada Musim Kemarau di Dusun Geliang Desa Pempatan. Buletin Udayana Mengabdi 17:92–96
6. Malalantang SS, Telleng MM, Kaunang WB, Waani MR, Sane S (2023) Silase Sebagai Salah Satu Solusi Mengatasi Kekurangan Hijauan Pakan di Desa Kasuratan Kecamatan Remboken Kabupaten Minahasa. Polbangtan 4
7. Gurawal I, Rawendra R (2022) Growth and nutrient content of corn fodder (Zea mays) by watering cow biourine. Jurnal Peternakan Indonesia 4:21–27
8. Liman L et al (2023) Effect of substitution of corn (Zea mays) fodder with sorghum (Sorghum bicolor var. Numbu) fodder on nutrient digestibility and daily weight gain of beef cattle. J Agricultural Sci Sri Lanka 18:329–336. https://doi.org/10.4038/JAS.V18I3.9449
9. Ru L., Shu-ying Z., D.M. and Y.Z.: Design of raindrop detector based on Arduino. IEEExplore. (2021).
10. Baradaran AA, Tavazoei MS (2022) Fuzzy system design for automatic irrigation of agricultural fields. Expert Syst Appl 210:118602. https://doi.org/10.1016/j.eswa.2022.118602
11. dos Reis JC et al (2023) Fuzzy logic indicators for the assessment of farming sustainability strategies in a tropical agricultural frontier. Agron Sustain Dev 43. https://doi.org/10.1007/s13593-022-00858-5
12. Yadav R, Daniel AK (2018) Fuzzy based smart farming using wireless sensor network. In: 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), pp 1–6. https://doi.org/10.1109/UPCON.2018.8597086
13. Jayalakshmi B, Haritha H, Maniyeri A, Arjun N (2020) Fuzzy based irrigation and lighting systems for indoor farming. In: 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), pp 1183–1186. https://doi.org/10.1109/ICIRCA48905.2020.9183154
14. Dela Cruz JR, Magsumbol J-AV, Dadios EP, Baldovino RG, Culibrina FB, Gan Lim LA (2017) Design of a fuzzy-based automated organic irrigation system for smart farm. In: 2017 IEEE 9th international conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), pp 1–6. https://doi.org/10.1109/HNICEM.2017.8269500
15. Haryanto, Ulum M, Ibadillah AF, Alfita R, Aji K, Rizkyandi R (2019) Smart aquaponic system based internet of things (IoT). J Phys.: Conference Series.
16. Madhuri A, Jaswanth K, Bhavana B, Shankar KB (2023) Auto indoor hydroponic fodder grow chamber using IoT. Int J Recent Adv Multidisc Topics 4:40–45
17. Raghavendran PS, Logeswaran T, Ambayeram M, Hemasree T, Keerthana S, Naveen J (2019) Microcontroller based efficient automated cattle fodder system. Int J Innov Technol Exploring Eng (IJITEE). 9
18. Blanco-Gómez P, Mateu-Belloch A, Jiménez-García JL, Salas-Cantarellas AJ, Pieras-Company JJ, Santamaría-Casals E (2024) Real-time ultrasonic water level IoT sensor for in-situ soil permeability testing. HardwareX 19. https://doi.org/10.1016/j.ohx.2024.e00541
19. Pozzebon A, Benini M, Bocci C, Fort A, Parrino S, Rapallo F (2023) Grid-layout ultrasonic LoRaWAN-based sensor networks for the measurement of the volume of granular materials. Measurement (Lond) 220:113404. https://doi.org/10.1016/j.measurement.2023.113404
20. Blümel K, Tagliaferri F, Kuhl M (2023) Algorithm for calculating distance and sensor-object angle from raw data of ultra-low power, long-range ultrasonic time-of-flight range sensors. Proc CIRP 118:1061–1065. https://doi.org/10.1016/j.procir.2023.06.182

21. Guerrero JSG, González AFC, Vega JIH, Tovar LAN (2015) Instrumentation of an Array of ultrasonic sensors and data processing for unmanned aerial vehicle (UAV) for teaching the application of the Kalman filter. Proc Comput Sci 75:375–380. https://doi.org/10.1016/j.procs.2015.12.260
22. Sahoo AK, Udgata SK (2024) Material classification based on non-contact ultrasonic Echo signal using deep learning approach. Proc Comput Sci 235:606–616. https://doi.org/10.1016/j.procs.2024.04.060
23. Kapalamula HE, Mlatho JS, Macheso PS (2022) Design and implementation of a social distance vest for Covid19 prevention (SODIV-COP). Int J Intell Netw 3:113–118. https://doi.org/10.1016/j.ijin.2022.08.003
24. Setya W et al (2019) Design and development of measurement of measuring light resistance using light dependent resistance (LDR) sensors. J Phys Conf Ser 1402. https://doi.org/10.1088/1742-6596/1402/4/044102
25. Ojeda CM, Dela Cruz JC (2020) Analysis on the optical response of Ldr and phototransistor in photoplethysmography. Sens Biosensing Res 28:100334. https://doi.org/10.1016/j.sbsr.2020.100334
26. Pradhan S, Nayak A, Sharma GA (2021) Smart lighting system using light dependent resistor [Ldr]. Int J Acad Appl Res 5:32–37
27. Al-Nahhas YS, Hadidi LA, Islam MS, Skitmore M, Abunada Z (2024) Modified Mamdani-fuzzy inference system for predicting the cost overrun of construction projects. Appl Soft Comput 151:1–23
28. Pandey SKS (2023) An optimal hybrid multiclass SVM for plant leaf disease detection using spatial fuzzy C-means model. Expert Syst Appl 214
29. Saade JJ, Diab HB, Defuzzification methods and new techniques for fuzzy controllers (2004) Iran J Electr Comput Eng 3:161–174
30. Cervero A, Castro-Lopez A, Álvarez-Blanco L, Esteban M, Bernardo A (2020) Evaluation of educational quality performance on virtual campuses using fuzzy inference systems. Eur J Psychol Educ 15
31. Thomas (2019) A novel fuzzy based assessment method for learning outcomes. Int J Recent Technol Eng (IJRTE) 8:1665–1669
32. Torres A, Nieto JJ (2006) Fuzzy logic in medicine and bioinformatics. J Biomed Biotechnol
33. De A, Singh SP (2021) Analysis of fuzzy applications in the agri-supply chain: a literature review. J Clean Prod 283
34. Shafaei SM, Nourmohamadi-Moghadami A, Kamgar S (2023) Application of adaptive neuro-fuzzy simulation system for elucidation of water absorption process of lentil seeds. Measurement: Food 10:100087. https://doi.org/10.1016/j.meafoo.2023.100087
35. Madhuri A, Jaswanth K, Bhavana B, Shankar KB (2023) Auto indoor hydroponic fodder grow chamber using IoT. Int J Recent Adv Multidisc Topics 4:40–45
36. Raghavendran PS, Logeswaran T, Ambayeram M, Hemasree T, Keerthana S, Naveen J (2019) Microcontroller based efficient automated cattle fodder system. Int J Innov Technol Exploring Eng (IJITEE) 9

A Monitoring System for Dissolved Oxygen and Temperature at Specific Water Depths Using Water Pressure Sensors

Muhammad Akmal Falih Rizqullah[1(✉)], Luthfi As-Sidqi[1], Mikhail Sannata Widjaya[1], Fiky Y. Suratman[2], Istiqomah[2], Eka Wardhani[3], and Kancitra Pharmawati[3]

[1] School of Electrical Engineering, Telkom University, Bandung, Indonesia
akmalfalih@student.telkomuniversity.ac.id
[2] The University Center of Excellence Intelligence Sensing-IoT, Telkom University, Bandung, Indonesia
[3] Environmental Engineering, ITENAS, Bandung, Indonesia

Abstract. Dissolved oxygen (DO) and temperature are key indicators of water quality, significantly affecting aquatic ecosystems. This study develops a monitoring system to measure DO and temperature at multiple water depths with high accuracy, utilizing water pressure sensors for precise depth determination. The system's design includes integrating DO and temperature sensors with water pressure sensors to monitor water quality at various depths, with data processed through custom software. Regression analysis is used to calibrate the sensors, ensuring accurate readings. The system's design includes a submerged unit for collecting water data and a floating unit for controlling depth and displaying the data. The system was tested for accuracy and effectiveness, showing an average error of ± 2.41 cm in depth measurement and an R^2 value of 0.9955 for DO sensor accuracy, indicating minimal deviation from actual values. Statistical analysis, including mean, regression, and correlation analyses, confirmed the system's precision, with significant variations in DO levels recorded at different depths, reaching a maximum of 6 mg/L at 140 cm. These results highlight the system's capability to provide reliable, real-time water quality data across various depths, proving its effectiveness for environmental monitoring.

Keywords: Water quality · Depth · Monitoring · Dissolved oxygen · Water pressure · Temperature

1 Introduction

Water is a vital component of life on Earth, playing an essential role in the survival of all living organisms, including humans. However, water pollution has become a pressing global issue, especially in Indonesia. According to the Ministry of Environment of the Republic of Indonesia, 75% of rivers in the country were polluted due to household waste disposal as of 2014. This alarming statistic underscores the urgent need for effective measures to prevent further degradation of water quality.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 305–318, 2026.
https://doi.org/10.1007/978-981-95-9233-3_23

Water quality monitoring extends beyond environmental conservation, playing a crucial role in various sectors such as aquaculture, where the well-being of farmed fish hinges on maintaining proper levels of dissolved oxygen and temperature [1]. In agriculture, water quality significantly influences crop irrigation, especially in regions facing water shortages or contamination [2]. Additionally, municipal water systems depend on constant monitoring to guarantee the provision of safe drinking water for city residents [3]. Climate research also utilizes water quality data to observe changes in aquatic ecosystems caused by global warming, including the rise in ocean acidification and thermal stratification [4].

To address this, various methods have been developed for water quality monitoring. Traditionally, water quality is assessed by collecting samples from the water surface and analyzing them in a laboratory [5]. Collecting and analyzing the samples in the laboratory provides highly reliable and detailed data on water quality. However, this method does not support sampling at a specific depth and is time-consuming, as it requires transportation and laboratory analysis, which do not reflect real-time changes in water quality. More advanced techniques involve using devices like the Kemmerer Bottle Sampler shown in Fig. 1 to collect water from specific depths [6]. Although this method can take samples at a certain depth, this method still requires manual operation and may not be feasible for continuous monitoring, limiting its ability to capture real-time changes in water quality at varying depths over time. Recent studies have also explored online monitoring systems for water quality, though these are typically limited to surface measurements [7–9]. While these methods provide valuable data, they are prone to sample contamination and may not accurately reflect conditions at different water depths.

Fig. 1. Kemmerer bottle sampler

While numerous water quality monitoring systems exist, many are limited to surface measurements or require costly and labor-intensive sampling procedures that are prone to contamination uno [10–13]. Some water quality monitoring systems have advantageous features, such as transmitting water quality data via GSM and IoT, but they are still limited to surface-level measurements only [14, 15]. Furthermore, existing systems often fail to account for depth-specific variations in water quality parameters, which

are crucial in environments like lakes and reservoirs that undergo thermal stratification. This study addresses this gap by developing a comprehensive monitoring system capable of providing accurate, real-time measurements of dissolved oxygen and temperature at various water depths using water pressure sensors for precise depth determination.

This paper is structured as follows: Sect. 2 discusses the design and development of the floating and submerged devices used in the monitoring system. Sect. 3 details the implementation of the system, including the deployment and data collection processes. Sect. 4 presents the results of the experiments conducted to validate the system's performance. Sect. 5 presents the explanation of the result, a comparison among other studies, weaknesses, and implications. Finally, Sect. 6 offers conclusions and suggestions for future work.

2 Materials and Method

2.1 Water Quality

Water quality is a measure that determines the condition of water based on its physical, chemical, and biological characteristics. This condition is measured to ensure that the water is suitable for various purposes, such as human consumption, agriculture, industry, and aquatic ecosystems. The main parameters that affect water quality include temperature, pH, dissolved oxygen (DO), the content of harmful chemicals such as heavy metals, and the amount of nutrients such as nitrogen and phosphorus [7]. However, this study focuses on dissolved oxygen and water temperature as water quality parameters.

Dissolved oxygen (DO) is one of the most critical water quality parameters. DO refers to the amount of free oxygen available in water and is a key indicator of the health condition of aquatic ecosystems. This oxygen is needed by aquatic life for respiration [8]. Low DO levels can cause hypoxic conditions that are harmful to aquatic life, while levels that are too high can cause supersaturation which is also detrimental. DO levels in water are influenced by various factors such as water temperature, salinity, atmospheric pressure, and biological and chemical activities in the water [16].

Water temperature is a crucial parameter in evaluating water quality as it affects several biological, chemical, and physical processes in aquatic ecosystems. Thus, water temperature can provide an idea of the thermal energy possessed by water molecules which can have a major impact on various physical and chemical phenomena, including solubility of substances, electrical conductivity, and the ability of water to store heat. Water temperature also affects the level of dissolved oxygen in a body of water [7].

The Indonesian government has provided standards for the classification of acidity levels in waters contained in Appendix VI of Government Regulation of the Republic of Indonesia Number 22 of 2021 concerning the Implementation of Environmental Protection and Management in the Quality Standards for Lake and Similar Waters section. Table 1 shows the classification of water quality against the parameters of water temperature and dissolved oxygen in water.

Table 1. Class of dissolved oxygen (mg/L) and temperature (°C) [17]

Class	Dissolved oxygen (mg/L)	Temperature (°C)
1	6	Dev 3
2	4	Dev 3
3	3	Dev 3
4	1	Dev 3

According to Table 1, lake water quality is categorized into four classes: class 1, class 2, class 3, and class 4. Class 1 water is suitable for drinking or other uses that require high-quality water. Class 2 water is used for facilities or infrastructure, water recreation, freshwater fish farming, animal husbandry, or similar purposes. Class 3 water is used for freshwater fish farming, animal husbandry, irrigation, or similar purposes. Class 4 water is used for irrigation or similar purposes. Therefore, as long as lake water meets the criteria for any of these classes, it is considered suitable for use based on its quality. However, if lake water does not meet these criteria, it is deemed unsuitable for use or contaminated.

Then, the class for the temperature parameter is divided based on dev 3 which means it has a deviation value of 3. This means that good quality lake water should have a standard deviation of 3 against the air temperature above the water surface. For example, if a lake has an air temperature above the water surface of 28 °C, then good quality lake water will have a temperature ranging from 25 °C to 31 °C. As explained earlier about water temperature, this can be influenced and affect the biological and chemical processes that occur in the water.

2.2 Depth Measurement

This research uses hydrostatic pressure to determine the depth position of submerged devices. Hydrostatic pressure is the pressure exerted by the weight of the fluid at a specific point due to gravity within a fluid column. According to Pascal's Law, fluid pressure at a given depth spreads evenly in all directions. Therefore, the hydrostatic pressure remains constant at the same depth from the water surface, providing an accurate measurement of depth position.

$$P_h = \rho\ g\ h \tag{1}$$

$$\rho = \frac{m}{V} \tag{2}$$

$$g = \frac{GM}{R^2} \tag{3}$$

Based on Eq. (1), hydrostatic pressure (Ph) is directly proportional to the density of the fluid (ρ), the acceleration of gravity (g), and the depth of the review point in the fluid

(h). The density of a fluid can be determined by comparing the mass to the volume of the fluid as in Eq. (2). The density of water commonly used in calculations is 1 g/cm^3 or 1000 Kg/m^3. Then, the acceleration of gravity on Earth can be calculated using Eq. (3) which depends on the gravitational constant (G), the mass of the Earth (M), and the distance between the review point and the center of mass (R) [18].

2.3 System Design

To achieve the objectives of this research, we developed a system consisting of two units: the Submerged unit and the Floating unit. The Submerged unit is responsible for collecting water quality data, such as temperature and dissolved oxygen levels. The Floating unit determines the depth position for data collection, controls the depth position of the Submerged unit, and receives data from the Submerged unit for user display via an LCD. This setup is illustrated in Fig. 2.

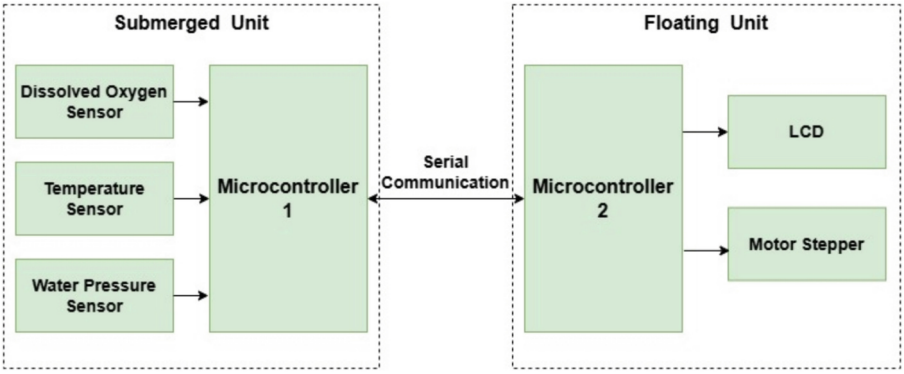


Fig. 2. Block diagram of the system

Figure 3 is the implementation of the system design for the Submerged Unit. To meet safety requirements, it must have an IP68 rating, that is water resistant in fresh water to a maximum depth of 1.5 m for up to 30 min.

2.4 Regression Analysis

Regression analysis is a statistical method used to model the relationship between a dependent variable and one or more independent variables [19]. In the context of water quality measurement, particularly for measuring parameters such as dissolved oxygen and water temperature at various depths, regression analysis can be used to calibrate the sensor by linking the water pressure data obtained from the sensor with the actual water depth. Equation (4) is the simple linear regression.

$$y = \beta_0 + \beta_1 x + \in \tag{4}$$

where y represents the dependent variable, which in our case could be the sensor measurement, such as dissolved oxygen or temperature. The independent variable, denoted

Fig. 3. Implementation of the submerged unit

by x, is a factor such as water pressure or depth. β_0 is the intercept, indicating the value of y when x is zero, while β_1 represents the slope of the regression line, illustrating the rate of change in y concerning x. The error term, $\in$, accounts for the discrepancies between the observed values and the predicted values from the regression model. In this study, simple linear regression is applied to analyze the relationship between water pressure and depth, allowing us to calibrate the water pressure sensor accurately. By fitting a regression line to the observed data, we can predict the depth based on pressure readings. A high R^2 value from the regression analysis indicates a strong correlation, suggesting that the water pressure sensor provides reliable depth measurements with minimal error. This calibration is essential for ensuring the accuracy of the system when monitoring dissolved oxygen and temperature at various water depths.

2.5 Sample Description and Data Acquisition

In this study, data were collected for dissolved oxygen (DO), temperature, and water pressure to measure depth. For each experiment, ten readings were collected per sample, and the average value was calculated. The samples included:

- **Dissolved Oxygen (DO)**: Four samples (0 mg/L, 1.6 mg/L, 4.5 mg/L, and 5.8 mg/L).
- **Temperature**: Four samples (20 °C, 25 °C, 37.5 °C, and 40 °C).
- **Water Pressure (Depth)**: Eight samples representing various depths.

The primary objective was to monitor water depths up to 4 m, DO levels up to 6 mg/L, and temperatures from 20 °C to 40 °C. The final integrated experiment was conducted in a swimming pool environment, where all measurements were taken. The data acquisition process involved two Arduino Uno microcontrollers to interface with the sensors. Calibration was conducted before data collection: the DO and temperature sensors were calibrated in the laboratory, while the water pressure sensor was calibrated in the swimming pool to ensure accuracy at the measurement site. For each sample, ten measurements were recorded and averaged to reduce variance, providing reliable data for analysis.

3 Result

Before assembling the system for this research, the sensors used need to be tested for performance first with the main parameter being accuracy. The analysis method used is regression [20, 21]. If the performance target has not been achieved, then sensor calibration is carried out by inserting the regression equation that has been obtained in the code in the microcontroller to improve performance.

3.1 Dissolved Oxygen Sensor

Dissolved oxygen sensor testing is done using a solution made to have certain dissolved oxygen levels. In this research, the solution used has dissolved oxygen levels of 0 mg/L, 1.6 mg/L, 4.5 mg/L, and 5.8 mg/L. Regression analysis on the DO sensor produces the equation shown in Fig. 4 with an R^2 score of 0.9955. This shows that the sensor reading has a very small reading error against the actual data. Thus, this sensor does not need to be recalibrated.

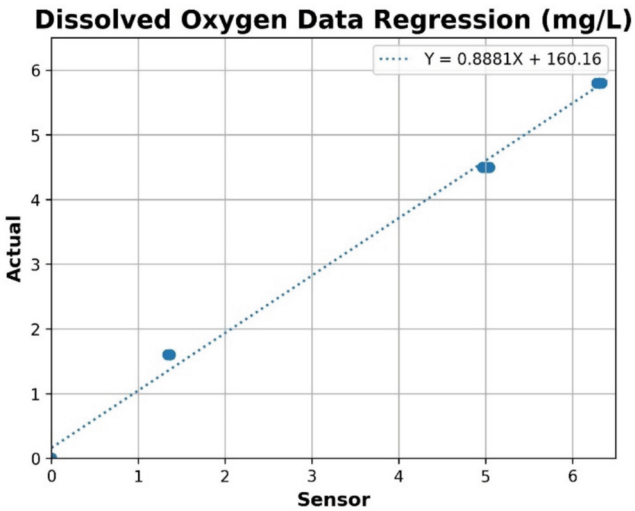


Fig. 4. Dissolved oxygen data regression (mg/L)

3.2 Temperature Sensor

Temperature sensor testing is done using a solution with varying temperatures, specifically 20 °C, 25 °C, 37.5 °C, and 40 °C. Then, the sensor reading results are compared with the reading results of standardized measuring instruments. Regression analysis on the temperature sensor produces the equation shown in Fig. 5 with an R^2 score of 1. This shows that the sensor reading has a very small reading error against the actual data which can be ignored. Thus, this sensor does not need to be recalibrated.

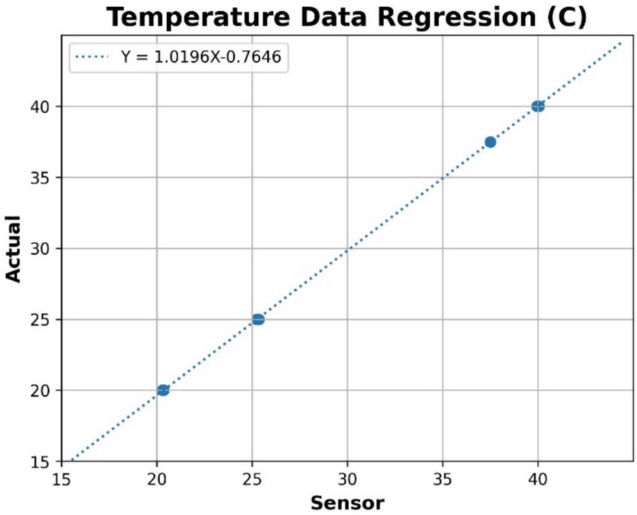


Fig. 5. Temperature data regression (C)

3.3 Water Pressure Sensor

The water pressure sensor testing is to measure depth which is done by submerging the sensor at a certain depth. By applying the principle of hydrostatic pressure, the equation of depth to water pressure can be entered into the microcontroller code. This test was conducted at a depth of 500 mm, 1000 mm, 1500 mm, 2000 mm, 2500 mm, 3000 mm, 3500 mm, and 3900 mm.

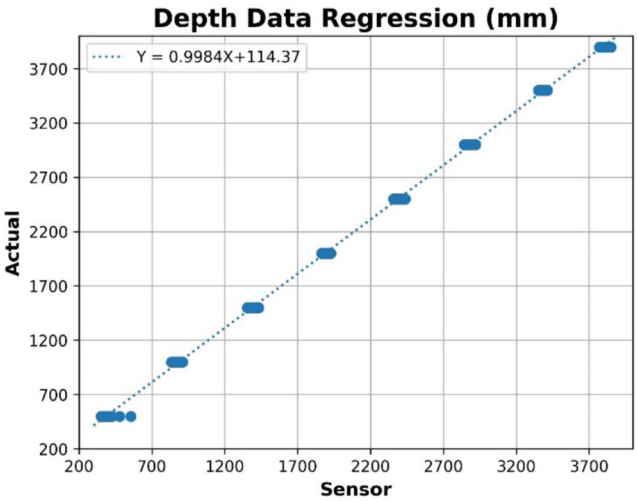


Fig. 6. Depth data regression (mm)

Regression analysis on the water pressure sensor is to determine depth which produces the equation shown in Fig. 6 with an R^2 score of 0.9992. This shows that the sensor reading has a small reading error against the actual data, but there are some outliers in the data. So, although the R^2 score is quite good, this sensor needs to be recalibrated so that the readings are more stable and reliable. After calibrating the water pressure sensor for depth measurement, tests were conducted to verify that the target performance was met. The testing process mirrored the calibration method, using various depth positions as reference points, as outlined in Table 2.

Table 2 Depth sensor error reading

Actual depth (cm)	Sensor reading (cm)	Error (cm)
50	51.4	1.4
100	99.8	0.2
150	149.1	0.9
200	198.0	2
250	249.5	0.5
300	296.7	3.3
350	341.4	8.6

Based on Table 2, the water pressure sensor generally provides accurate readings with an average error of 2.41 cm and a standard deviation of 2.70 cm. However, there is a notable increase in error at greater depths, specifically at 350 cm. This indicates that while the sensor performs well at shallower depths, its accuracy decreases with increasing depth, necessitating potential recalibration or adjustment for deeper measurements

3.4 Integration

After all the sensors met their performance targets, the entire system was tested by submerging it at various specific depths to collect water quality data. In this test, the system will be submerged to a depth of up to 140 cm.

Figure 7 shows the correlation between dissolved oxygen (DO) and depth, indicating a linear increase in DO levels with increasing depth, with a notable maximum DO level of 6 mg/L observed at depths of 120 cm and 140 cm. This trend suggests that as measurement depth increases, the DO concentration in the water also rises, reflecting a potential pattern in oxygen distribution. These findings align with those reported in Tan et al.'s study [22], which demonstrated a strong inverse correlation between water temperature and DO levels. In both studies, it was observed that lower water temperatures at greater depths correlate with higher dissolved oxygen concentrations, reinforcing the understanding that colder water holds more oxygen.

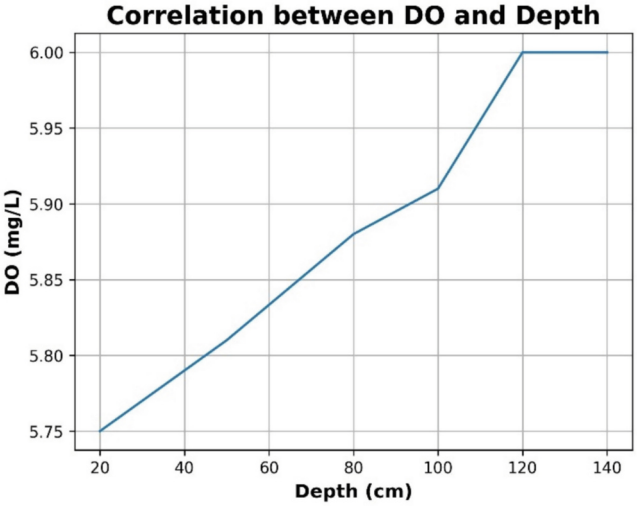


Fig. 7. Correlation between DO and depth

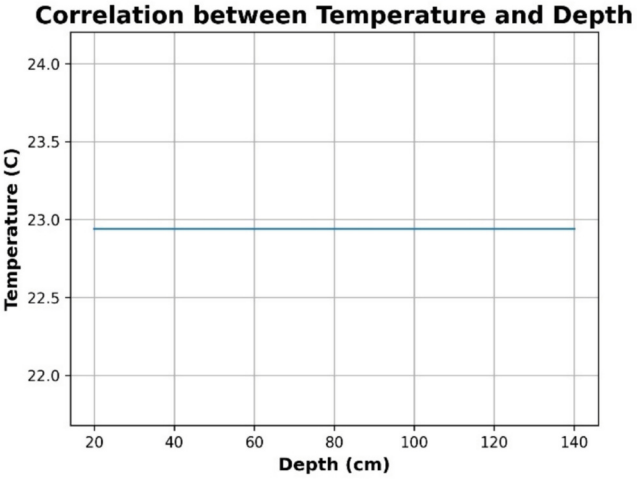


Fig. 8. Correlation between temperature and depth

In contrast, based on Figs. 8 and 9, the temperature data reveals no significant change with depth. The temperature remains constant at 22.94 °C at each measured depth, which indicates that the water temperature is uniformly distributed throughout the depth range tested. This uniformity in temperature across different depths precludes the possibility of establishing a correlation between DO levels and temperature, as there is no variation in the temperature data to compare against the DO levels.

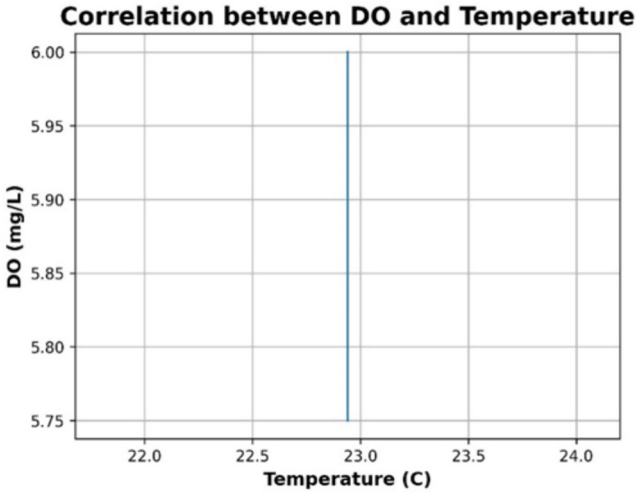


Fig. 9. Correlation between DO and temperature

The observed phenomena imply that the water flow at the test site is highly effective, ensuring a relatively even distribution of DO levels throughout the water column. The slight increase in DO levels with depth could be due to various factors, such as the presence of additional oxygen sources at greater depths. One plausible explanation is the occurrence of biological activities, such as photosynthesis by aquatic plants or algae, which produce oxygen and contribute to the increased DO levels observed at these depths [23]. Furthermore, the effective water flow could facilitate the mixing of water layers, preventing the stratification that often leads to varying DO levels at different depths [24]. This mixing process ensures that oxygen produced at the surface or various depths is evenly distributed throughout the water column, resulting in the relatively uniform DO levels recorded in the study.

4 Discussion

The results of our study indicate that the dissolved oxygen (DO) sensor performed well, showing minimal deviation from actual DO levels, while the temperature sensor consistently recorded uniform readings across various depths. This outcome suggests that the system is capable of providing reliable, real-time data on water quality parameters. Notably, the observed increase in DO concentration at greater depths could be attributed to oxygen-producing biological processes, which may be more active in deeper water layers. Studies such as those by Patty confirm that increased photosynthetic activity by aquatic plants or algae at greater depths contributes to higher oxygen levels [16]. This correlation between depth and DO supports our initial hypothesis that the integration of water pressure sensors to monitor specific depths would yield accurate insights into water quality variations.

When comparing our results with previous studies, we observed that surface-only monitoring systems often miss subsurface variations in DO and temperature [25]. For

instance, while systems such as those described by Quevy et al. provide useful surface data, they lack the depth-specific differentiation that our system captures, offering a more complete understanding of water quality dynamics for environmental modeling and interpretation [11, 26]. Similar findings were reported in studies utilizing the Kemmerer Bottle Sampler for depth-based water sampling, although these traditional methods do not offer the continuous, real-time data provided by our system.

However, our study does present some weaknesses. The water pressure sensor showed reduced accuracy at greater depths, with increasing error beyond 300 cm, similar to findings from Stenvold regarding water pressure measurement errors in deep-water environments [21]. Although recalibration improved performance, further refinement is necessary for consistent accuracy at greater depths. Additionally, like previous studies, our system focused on DO and temperature but omitted other critical parameters such as pH, turbidity, and nutrient levels, which are essential for a more holistic assessment of water quality [5, 27–29]. The implications of this study for water quality management are significant, particularly in environments where stratification and depth-related variations are critical [30]. By providing real-time, depth-specific data, our system could improve monitoring efforts, enabling more informed decision-making related to pollution control and ecosystem management. The system's potential for IoT integration also opens up opportunities for large-scale environmental monitoring [9, 11, 31–33].

5 Conclusion

The DO and temperature sensors perform well, with R^2 scores close to 1, indicating no need for recalibration. Although the water pressure sensor also has a good R^2 score, recalibration is required as this research focuses on using the water pressure sensor to measure water quality parameters at depth. Based on the test results, the water pressure sensor performs well for monitoring water quality at depth, with an average error of ± 2.41 cm. Future research could expand the measured water quality parameters to include pH, electrical conductivity, and turbidity. Additionally, development could incorporate multi-node features using IoT.

Acknowledgments. The research with the title "A Monitoring System for Dissolved Oxygen and Temperature at Specific Water Depths Using Water Pressure Sensors" received a lot of support in the form of funding from the IoT Research Center and Telkom University.

References

1. Salin KR, Arome Ataguba G (2018) Aquaculture and the environment: towards sustainability. In: Sustainable Aquaculture. Springer International Publishing, Cham, pp 1–62. https://doi.org/10.1007/978-3-319-73257-2_1
2. Matteoli F, Schnetzer J, Jacobs H (2020) Climate-smart agriculture (CSA): an integrated approach for climate change Management in the Agriculture Sector. In: Handbook of climate change management. Springer International Publishing, Cham, pp 1–29. https://doi.org/10.1007/978-3-030-22759-3_148-1

3. Peters-Lidard CD et al (2021) Indicators of climate change impacts on the water cycle and water management. Clim Chang 165:36. https://doi.org/10.1007/s10584-021-03057-5
4. Yadav NK et al (2024) Climate change effects on aquaculture production and its sustainable management through climate-resilient adaptation strategies: a review. Environ Sci Pollut Res 31:31731–31751. https://doi.org/10.1007/s11356-024-33397-5
5. Onyango J, Kitaka N, van Bruggen JJA, Irvine K, Simaika J (2024) Agricultural intensification in Lake Naivasha catchment in Kenya and associated nutrients and pesticides pollution. Sci Rep 14:18539. https://doi.org/10.1038/s41598-024-67460-5
6. Endah Purnamaningtyas S (2014) Distribusi Konsentrasi Oksigen, Nitrogen, dan Fosfat di Waduk Saguling, Jawa Barat
7. Elfidasari D, Noriko N, Effendi Y, Puspitasari RL (2015) Water quality of situ Lebak Wangi Bogor. Al-Azhar Indonesia Seri Sains dan Teknologi 3
8. Wahyuningsih S, Dharmawan A (2020) Determination of reaeration coefficient of Bedadung Hilir River oxygen deficit change method (case study in Balung District, Jember). Presipitasi 17:169–176
9. Yaacob N, Zainali NS, Rahman AAA, Yusof AL, Kassim M, Salehudin ASN (2025) Design of Water Quality Monitoring System Based on internet of things technology. J Adv Res Appl Sci Eng Technol 45:154–167. https://doi.org/10.37934/araset.45.1.154167
10. Wahjono HD (2015) Water quality monitoring for Semayang Lake and Melintang Lake in Kutai Kartanegara District using online monitoring system
11. Quevy Q, Lamrini M, Chkouri M, Cornetta G, Touhafi A, Campo A (2023) Open sensing system for long term, low cost water quality monitoring. IEEE Open Journal of the Industrial Electronics Society 4:27–41. https://doi.org/10.1109/OJIES.2022.3233919
12. Gleiser M, Moro S (2023) Implementation of an IoT-based water quality monitoring system for aquaculture. Int J Res Publ Rev 4:1449–1452. https://doi.org/10.55248/gengpi.234.5.38043
13. Hakimi IM, Jamil Z (2021) Development of water quality monitoring device using Arduino UNO. IOP Conf Ser Mater Sci Eng 1144:012064. https://doi.org/10.1088/1757-899x/1144/1/012064
14. Demetillo AT, Japitana MV, Taboada EB (2019) A system for monitoring water quality in a large aquatic area using wireless sensor network technology. Sustain Environ Res 1. https://doi.org/10.1186/s42834-019-0009-4
15. Lakshmikantha V, Hiriyannagowda A, Manjunath A, Patted A, Basavaiah J, Anthony AA (2021) IoT based smart water quality monitoring system. Glob Trans Proceed 2:181–186. https://doi.org/10.1016/j.gltp.2021.08.062
16. Patty S (2015) Characteristics of phosphate, nitrate, and dissolved oxygen in Lembeh Strait waters, North Sulawesi. J Pesisir dan Laut Tropis 2
17. Indonesia (2021) Appendix VI of the Government Regulation of the Republic of Indonesia Number 22 of 2021 on the Implementation of Environmental Protection and Management regarding Water Quality Standards for Lakes and Similar Water Bodies
18. Halliday D, Resnick R, Walker J (2011) Fundamentals of physics. Wiley, Hoboken
19. Montgomery DC (2012) Introduction to linear regression. John Wiley & Sons, Inc, Canada
20. Hines JW, Gribok AV, Attieh I, Uhrig RE (2000) Improved methods for on-line sensor calibration verification. In: Proceedings of the 8 the international conference on nuclear engineering
21. Stenvold T, Eiken O, Zumberge MA, Sasagawa GS, Nooner SL (2006) High-precision relative depth and subsidence mapping from seafloor water-pressure measurements. SPE J 11:380–389. https://doi.org/10.2118/97752-PA
22. Tan W et al (2022) Correlation analysis of water temperature and dissolved oxygen based on water quality monitoring data. Presented at the. https://doi.org/10.1007/978-3-031-06761-7_28

23. Susanto JP, Riyadi A (2021) Garno. Y.S, The Suitability of Lake Toba's water in North Tapanuli Regency for Raw water and Recreational Water
24. Egeten, E.E., Makahinda, T., Tumimomor, F.: Vertical Profile and Area of Daily Dynamics of Water Temperature Gradient in Lake Tondano at Fish Farming and Non-Fish Farming Locations.
25. Cao H, Guo Z, Wang S, Cheng H, Zhan C (2020) Intelligentwide-area water quality monitoring and analysis system exploiting unmanned surface vehicles and ensemble learning. Water (Switzerland) 12. https://doi.org/10.3390/w12030681
26. Simeonov V, Einax J, Stanimirova I, Kraft J (2002) Environmetric modeling and interpretation of river water monitoring data. Anal Bioanal Chem 374:898–905. https://doi.org/10.1007/s00216-002-1559-5
27. Yau SR, Jalani J, Sadun AS, Rejab SM, John J (2025) Development of an aquaponics farming technology system using Arduino based on internet of things. J Adv Res Appl Sci Eng Technol 45:11–24. https://doi.org/10.37934/araset.45.2.1124
28. Herlambang A (2006) Water Pollut Mitigation Strategies JAI 2
29. Ogilvie A, Belaud G, Massuel S, Mulligan M, Le Goulven P, Calvez R (2018) Surface water monitoring in small water bodies: potential and limits of multi-sensor Landsat time series. Hydrol Earth Syst Sci 22:4349–4380. https://doi.org/10.5194/hess-22-4349-2018
30. Endah Purnamaningtyas S (2014) Distribution of oxygen, nitrogen, and phosphate concentrations in Saguling reservoir, West Java
31. Komarudin M, Septama HD, Yulianti T, Yudamson A, Hendri J, Arafat MAD (2021) Multi node sensors for water quality monitoring towards precision aquaculture. IOP Conf Ser Earth Environ Sci 739:012026. https://doi.org/10.1088/1755-1315/739/1/012026
32. Wiranto G, Mambu GA, Hiskia H, Widodo S (2015) Design of online data measurement and automatic sampling system for continuous water quality monitoring. In: 2015 IEEE international conference on mechatronics and automation (ICMA). IEEE, pp 2331–2335. https://doi.org/10.1109/ICMA.2015.7237850
33. Pule M, Yahya A, Chuma J (2017) Wireless sensor networks: a survey on monitoring water quality. J Appl Res Technol 15:562–570. https://doi.org/10.1016/j.jart.2017.07.004

Development of a Bedside Monitor with an Embedded Digital Filter System to Reduce Noise in the Electrocardiogram Signal

Bambang Guruh Irianto, Anita Miftahul Maghfiroh(✉), Reno Adi Surya Alsari, Abd. Kholiq, and Syevana Dita Musvika

Department of Electro-medical Technology, Politeknik Kesehatan Kemenkes Surabaya, Surabaya, Indonesia
anitamiftah@poltekkes-surabaya.ac.id

Abstract. Bedside monitors are vital in medical settings such as Intensive Care Units (ICU) and operating rooms, enabling continuous real-time monitoring of critical patient parameters like heart rate, blood pressure, and oxygen saturation. Among these, the Electrocardiogram (ECG) signal is particularly crucial for diagnosing cardiovascular diseases (CAD) by depicting the heart's electrical activity. However, the accuracy of ECG signals is often compromised by noise and interference from sources such as electrical disturbances and body movements. This research aims to enhance ECG signal clarity by employing Finite Impulse Response (FIR) digital filters to minimize noise. The research contributes to the field by providing a comparative analysis of filtering methods, with the Kaiser window method emerging as the most effective for reducing ECG signal noise. The data used is obtained from human body signals in the Lead II position. The data will then be processed and analyzed offline using digital filters. Three digital filtering techniques—Kaiser window, rectangular, and Tukey—were implemented and evaluated based on their Signal-to-Noise Ratio (SNR). The Kaiser window method demonstrated superior performance, achieving the highest SNR value of 8.49 dB, resulting in a cleaner, interference-free ECG signal display. The study's findings underscore the importance of advanced filtering techniques in improving the accuracy of ECG monitoring, which could lead to more reliable diagnoses and better patient outcomes. This has significant implications for clinical practices, as it enhances the reliability and clarity of critical patient data, ultimately supporting better clinical decision-making.

Keywords: Electrocardiogram · Digital Filter · Finite Impulse Response · Noise · Signal to Noise Ratio

1 Introduction

Bedside monitors are essential tools in the medical world, especially in intensive care units (ICU) and operating rooms [1]. They allow real-time monitoring of a patient's condition, providing vital information such as heart rate, blood pressure, and oxygen

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 319–335, 2026.
https://doi.org/10.1007/978-981-95-9233-3_24

saturation [2]. One of the most crucial components of a bedside monitor is the Electrocardiogram (ECG) signal, which provides a direct picture of the heart's electrical activity [3, 4]. However, a major challenge in ECG signal capture is the presence of noise or interference that can affect the accuracy and reliability of the data generated. ECG signals obtained from patients are often contaminated by various types of noise, such as electromagnetic interference, muscle movement, and electrode movement [5]. This noise can result in significant signal distortion, reduce the accuracy of clinical interpretation, and potentially result in an incorrect diagnosis [6]. Therefore, an effective method is needed to minimize this noise so that the resulting ECG signal is cleaner and more representative of the patient's real condition [7]. The use of digital filters is one solution that is widely applied to reduce noise in ECG signals. Digital filters can filter out certain frequency components of unwanted signals, such as noise, while retaining relevant information [8]. The development of bedside monitors equipped with specialized digital filters can improve the quality of the displayed ECG signal, thus supporting more accurate medical decisions [9]. Digital filters also allow customization according to specific noise characteristics, providing flexibility and efficiency in various clinical conditions [10]. Various types of digital filters have been used to correct this problem. However, it is difficult to utilize filters with fixed coefficients to eliminate some types of artifacts, as ECG signals are known to be non-stationary signals [11].

In 2006, a study was conducted Electrocardiogram (ECG) is widely used for the diagnosis of heart disease. In this paper, we propose a new ECG enhancement method based on the latest method of developed empirical mode decomposition (EMD). The EMD method is proven to be effective in removing high-frequency noise and baseline wander (BW) in ECG signals, including for real and synthetic noise, resulting in a cleaner and easier-to-interpret signal. The drawback of this study is that using the EMD method requires a lot of computational time to process the ECG signal, especially when separating the various frequency components in the signal. When the noise in the signal is low or simple, EMD can remove too many important signal parts, which can reduce the quality of the resultant signal [12].

In 2018, research was conducted to evaluate the performance of FIR and IIR filters using the least-square algorithm, a widely utilized method in noise cancellation and real-time systems due to its rapid convergence and high accuracy. The study found that FIR filters are more effective in suppressing high-frequency components compared to IIR filters. Additionally, while the standard deviation of IIR filters was closer to the desired output, FIR filters exhibited a lower standard deviation when dealing with high-frequency noise. Overall, the FIR filters yielded more precise results than the IIR filters under the proposed algorithm [13]. In 2021, another study focused on designing window-based FIR filters for noise removal in ECG signals, employing Rectangular, Kaiser, Hamming, and Bartlett Windows. The Signal-to-Noise Ratio (SNR) values were computed and compared across these methods, with the Kaiser Window achieving the highest SNR value of 0.0064 [14]. A 2016 study addressed the removal of high-frequency electromyograph (EMG) signal noise from ECG signals using an FIR low pass filter designed with Distributed Arithmetic (DA) architecture. The Kaiser Window-based low-pass FIR filter excelled at eliminating high-frequency EMG interference from noisy ECG signals with low-order filters. The denoising performance was assessed by calculating the SNR

and Mean Square Error (MSE), yielding an SNR of 0.6438 dB and an MSE of 0.0501. This performance surpassed that of Rectangular, Hanning, Hamming, Blackman, and Bartlett Windows [15]. Further research in 2015 examined ECG noise removal using both FIR and IIR digital filters through various methods. For FIR filters, Rectangular, Hann, Hamming, Blackman, and Kaiser Windows were applied, while IIR filters included Butterworth, Chebyshev I, Chebyshev II, and Elliptic types. The evaluation based on the waveform, power spectrum density (PSD), SNR, and MSE revealed that the Kaiser window of order 56 outperformed other FIR and IIR filters (orders 1, 2, and 3) in SNR and MSE, consistently demonstrating superior performance across different data sets [16]. With the rapid advancement of wearable and wireless technologies, real-time and routine ECG monitoring has garnered increasing attention, propelled by the widespread adoption of medical health technology [17–19]. Numerous articles and research efforts have emerged in this field, covering a range of topics including the filtering of ECG signals, extraction of QRS complexes and other waves, compression of ECG signals, automatic ECG signal processing methods, and live heart condition monitoring systems [20–24]. ECG signal interference commonly arises from physiological artifacts, high-frequency noise, and power line interference, which can lead to inaccurate ECG signal diagnoses [25–28]. Various digital filters have been employed to address these issues. However, due to the non-stationary nature of ECG signals, it remains challenging to effectively use filters with fixed coefficients to eliminate certain types of artifacts [12, 29–35]. The contribution of this paper is:

1. With this development, it is expected that bedside monitors can produce cleaner ECG signals, improve monitoring accuracy, and ultimately, provide significant benefits in patient care.
2. Improved quality and reliability of ECG data generated by the bedside monitor, which can improve the quality of diagnosis, reduce the possibility of medical errors, and increase the efficiency and effectiveness of patient care, especially in critical situations.
3. This study will also evaluate the effectiveness of the digital filter used, both in terms of the quality of the signal produced and from the perspective of clinical use in the field.

2 Materials and Method

2.1 Experimental Setup

This study involved respondents aged between 20 and 22 years with normal heart conditions. The number of respondents in this study was 10 people to determine the results of the electrocardiogram signal of each respondent. Data collection is carried out when the respondent is at rest. Tapping on the human body using lead ll, then analyzing the best digital filter to reduce noise in the electrocardiogram signal. Electrocardiogram signal measurement analysis using the human body and a phantom. Respondent data was taken and filtered using an analog filter, then digitally filtered to reduce interference signals originating from the electrical network. Data was taken from the respondent's body signals using ECG electrode cables with a sample frequency of 200 samples per second and signal analysis was carried out using the Matlab application. The signal is tapped using

electrode wires and processed using Arduino Nano to be filtered. The filtered signal will be displayed on the monitor screen using serial communication between Arduino and Raspberry.

Conducting research involving respondents requires careful ethical considerations and compliance with ethical guidelines and regulations. Our research has passed ethical testing and obtained a certificate of ethical approval with certificate number 206/HRECC.FODM/II/2023 issued by the Health Research Ethics Clearance Commission of the Faculty of Dentistry, Universitas Airlangga. This ethical approval ensures that the study has been reviewed and approved by the ethics committee or institutional review board to protect the rights, safety, and welfare of the participants. Consent to parents or legal guardians. The consent process outlines the nature of the study, potential risks and benefits, and how respondents' data will be used. The captured data will be analyzed using the FFT (Fast Fourier Transform) method and filtered using FIR digital filters using the Kaiser, rectangular, and Tukey window methods. From several filter methods, the SNR (Signal to Noise Ratio) value will be analyzed and the filter with the highest value will be entered into the system.

2.2 Materials and Tools

This study uses an analog circuit consisting of the basic instrument, bandpass filter, notch filter, noninverting amplifier, and adder. The frequency is tapped with a range of 0.1–102 Hz to get the ECG signal. Then use a notch filter to eliminate the frequency of 50 Hz. After the heart signal is tapped then the data is read by the Arduino and processed by a digital filter to get signal results that are clean from noise. After the ECG signal results are obtained, they are sent to the Raspberry to be displayed on the monitor. The digital filtering process is obtained using Eq. (1) as follows:

$$y = b[0] * x0 + b[1] * x1 + b[2] * x2 + b[3] * x3 \ldots \tag{1}$$

where y represents the result of the process, then b is the value of the filter coefficient, and x is the result of shifting the process of multiplying the coefficient by the previous input. To determine the best filter method that can reduce the noise the most, the SNR (Signal to Noise Ratio) value was analyzed. That is the ratio between the desired signal and the ratio of interfering signals. Eq. (2) is written as follows:

$$\text{SNR} = 10 \text{ Log } 10 \left(\frac{S}{N} \right). \tag{2}$$

where 'S' is the power of the desired signal and 'N' is the power of the unwanted noise. The SNR value is measured in decibels (dB), with a higheSNR value indicating that the desired signal is significantly stronger than the noise. This indicates that the signal quality is good and the information received is clearer. Whereas a low SNR value indicates that the desired signal is almost comparable to or even weaker than the noise signal. This indicates that the signal quality is poor and less information is obtained. Then the formula for knowing the value of the heart rate uses Eq. (3) as follows:

$$HR = \frac{sampling\ \ rate \times 60}{R - R\ interval} \tag{3}$$

where HR refers to Heart Rate. In this study, the module was evaluated by comparing its measurement results with those obtained from a standard patient monitor. The ECG signals used for testing were taken from lead II, with subjects consisting of infants aged between 1 day and 3 months..

2.3 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. Heart signals are tapped using electrode wires and enter the analog circuit. The circuit consists of a basic instrument, a bandpass filter, a notch filter, a non-inverting amplifier, and an adder. Using a bandpass filter frequency range of 0.1–102 Hz, then a notch filter to eliminate the frequency of 50 Hz.

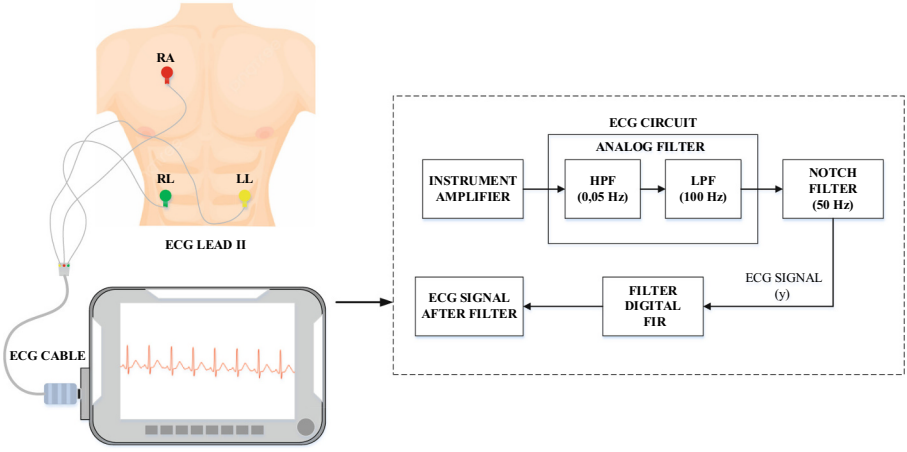


Fig. 1. Block diagram of bedside monitor with ECG circuit

From Fig. 1, it is explained that the input part has an ECG analog filter circuit to tap the heart signal. The basic instrument circuit functions to tap the heart signal that is tapped using electrodes. Then the signal will be filtered using a band pass filter with a frequency range of 0.05–100 Hz to get the ECG signal. Then a notch filter circuit is also needed to eliminate signals with a frequency of 50 Hz. Then the signal that appears will be amplified by an amplifier circuit so that it can be read and easily processed by the microcontroller. Then the microcontroller is useful as a process for processing ECG analog circuit data which will be converted into digital. In this part of the microcontroller, a digital filter will be embedded to produce ECG signals that are cleaner and free from noise signals. In addition to the microcontroller, in the process section, there is also a Raspberry that functions to receive and display data that has been processed by the microcontroller to be displayed on the monitor screen. The communication used between Arduino and Raspberry uses serial communication. Then the output section consists of an LCD screen that functions to display the parameters that have been processed. After

being received by Raspberry, the data will be displayed in the form of parameters listed on the LCD monitor screen.

The figure above illustrates the process of capturing body signals until the ECG signals are displayed on the LCD monitor. Initially, signals are obtained from electrodes attached to the respondent using ECG II leads. The data is then processed through various analog circuits such as amplifiers, bandpass filters, notch filters, and enhancers to produce the desired ECG signal while minimizing noise or unwanted signals. Next, the ECG signal is fed into the microcontroller, where an FIR digital filter is applied to further remove the noise. Finally, the processed ECG signal is displayed on a 10-in. LCD screen.

3 Result

After taking the respondent's body signal, the data is then processed using the Matlab application for offline filtering to determine the processing results by 3 types of digital filters, namely the Kaiser window, rectangular window, and Tukey window. **T**hen the signal results before and after the filter process of the three methods are analyzed comparatively. Figure 2(a) comparison of signals before and after filtering with the Kaiser window method, Fig. 2(b) comparison of signals before and after filtering with the rectangular window method, Fig. 2(c) comparison of signals before and after filtering with the Kaiser Tukey method.

In Fig. 2, the digital FIR Kaiser window filter was designed as a low-pass filter with a cutoff frequency of 40 Hz and a sampling rate of 200 samples per second. Then the raw ECG signal data was entered into the Matlab filter offline, and then the signal results were analyzed after filtering. Then the data is plotted in a graph, the blue signal represents the data obtained from body leads, while the orange signal displays the filtered ECG signal. From the analysis of the two signal images, it is found that the noise-filtered signal can be reduced so that the ECG signal can be seen more clearly than the signal before filtering. Fig. 2(b) - Digital FIR rectangular window filter design was created as a low pass filter with a cutoff frequency of 40 Hz and a sampling rate of 200 samples per second. Then the raw ECG signal data was entered into the Matlab filter offline and then the signal results were analyzed after filtering. The data is then plotted in the form of a graph, blue signal represents the data obtained from the body, while the orange signal displays the filtered ECG signal. From the analysis of the two signal images, it is found that the filtered signal looks clean and the noise can be reduced so that the ECG signal can be seen compared to the signal before filtering. Fig. 2(c) - Digital FIR Tukey window filter design utilizes a low pass filter with a cutoff frequency of 40 Hz and a sampling rate of 200 samples per second. Then the raw ECG signal data was entered into the Matlab filter offline and then the signal results were analyzed after filtering. The blue signal represents the data obtained from the body, while the orange signal displays the filtered ECG signal. From the analysis of the two signal images, it is found that the filtered signal looks clean and the noise can be reduced so that the ECG signal can be seen compared to the signal before filtering. However, when compared to the rectangular and Kaiser window methods, this method still has noise interference visible in the signal after filtering. After collecting data from the human body leads, the data is plotted in graph form for analysis. This

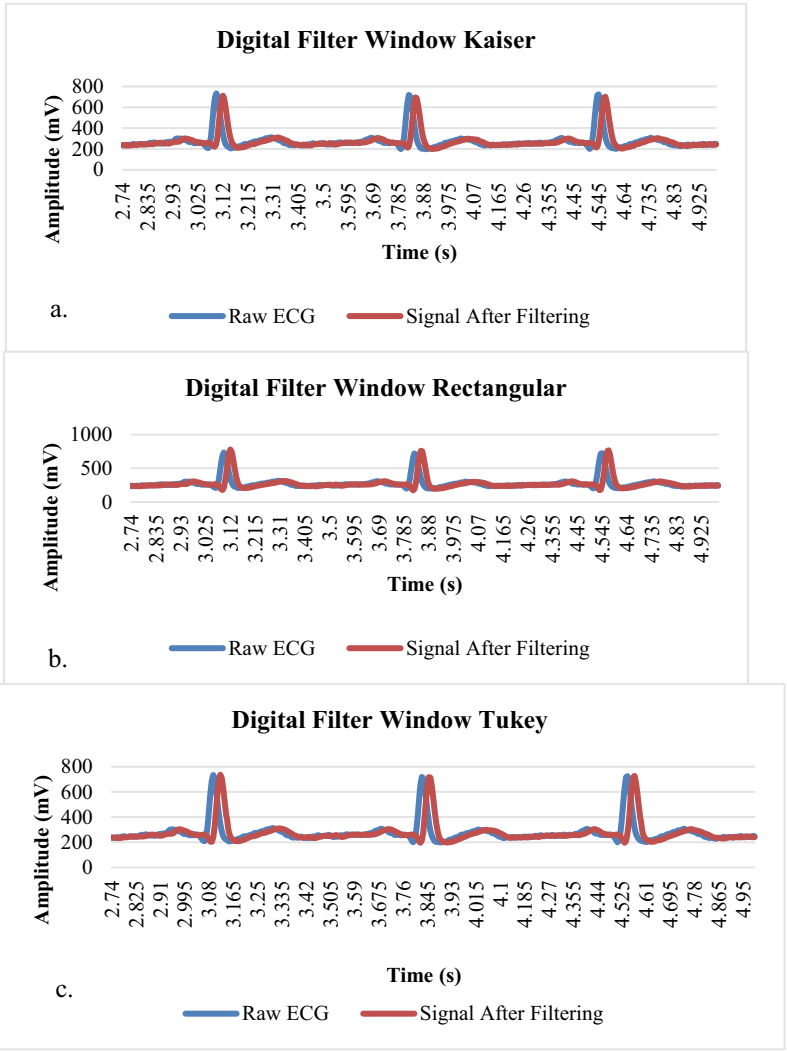
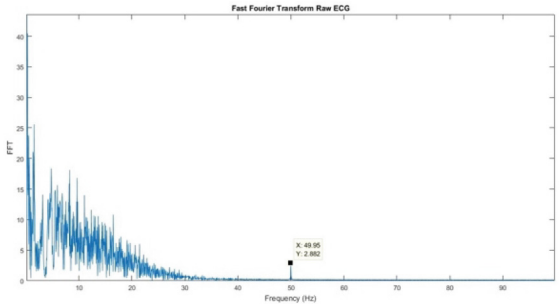
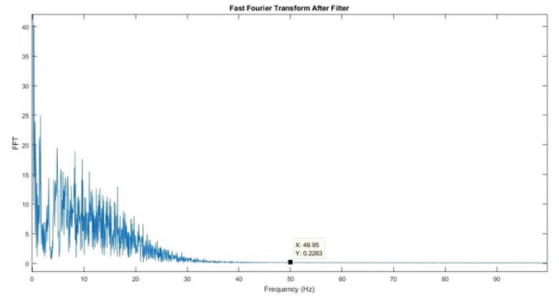


Fig. 2. Comparison of digital filter performance (a) Kaiser window, (b) rectangular window Fig. (c) Tukey window

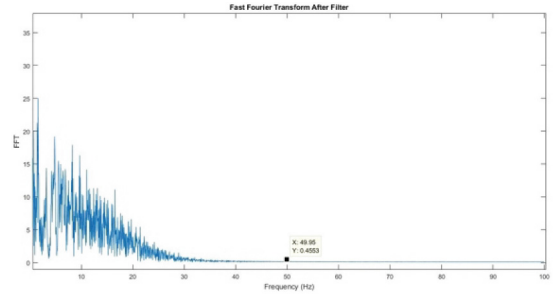
signal will be analyzed using FFT (Fast Fourier Transform) to identify the frequency of any interfering signals. FFT is an efficient algorithm for computing the discrete Fourier transform (DFT) and its inverse. The Fourier transform is a technique for converting data from the time domain to the frequency domain, which is highly beneficial for signal analysis and data processing. Figure 3(a) FFT of the raw signal, Fig. 3(b) FFT of Kaiser window method, Fig. 3(c) FFT of rectangular window method, Fig. 3(d) FFT of Tukey window method.



(a)



(c)



(d)

Fig. 3. The comparison of FFT signal among different types of filters (a) FFT of raw signal, (b) FFT of Kaiser window method, (c) FFT of rectangular window method, (d) FFT Tukey window method

In Fig. 3, FFT results from a raw signal with a dominance value of 59 Hz frequency signal is 2.882. Figure 3(b) - FFT (Fast Fourier Transform) result of the Kaiser digital window filter method. The dominance value of the 50 Hz frequency signal is 0.2453, down from before the filter process. Figure 3(c) - FFT (Fast Fourier Transform) rectangular window digital filter method. The dominance value of the 50 Hz frequency signal is 0.2263, down from before the filter process. Figure 3(d) - FFT (Fast Fourier Transform) of the Tukey window digital filter method. The dominance value of the 50 Hz

frequency signal is 0.4563, down from before the filter process. SNR (Signal to Noise Ratio) analysis is conducted to assess the ratio between the desired signal and the noise signal. This value is expressed in decibels (dB), where a higher SNR indicates that the desired signal is significantly stronger than the noise. This suggests good signal quality and clearer received information. Conversely, a lower SNR suggests that the desired signal is nearly comparable to or weaker than the noise signal. This indicates that the signal quality is poor and less information is obtained. The analysis is performed on each filter method and depicted in graphical form. Figure 4(a) SNR graph of the raw signal, Fig. 4(b) SNR graph of the Kaiser window method, Fig. 4(c) SNR graph of the rectangular window method, Fig. 4(d) SNR graph of the Tukey window method.

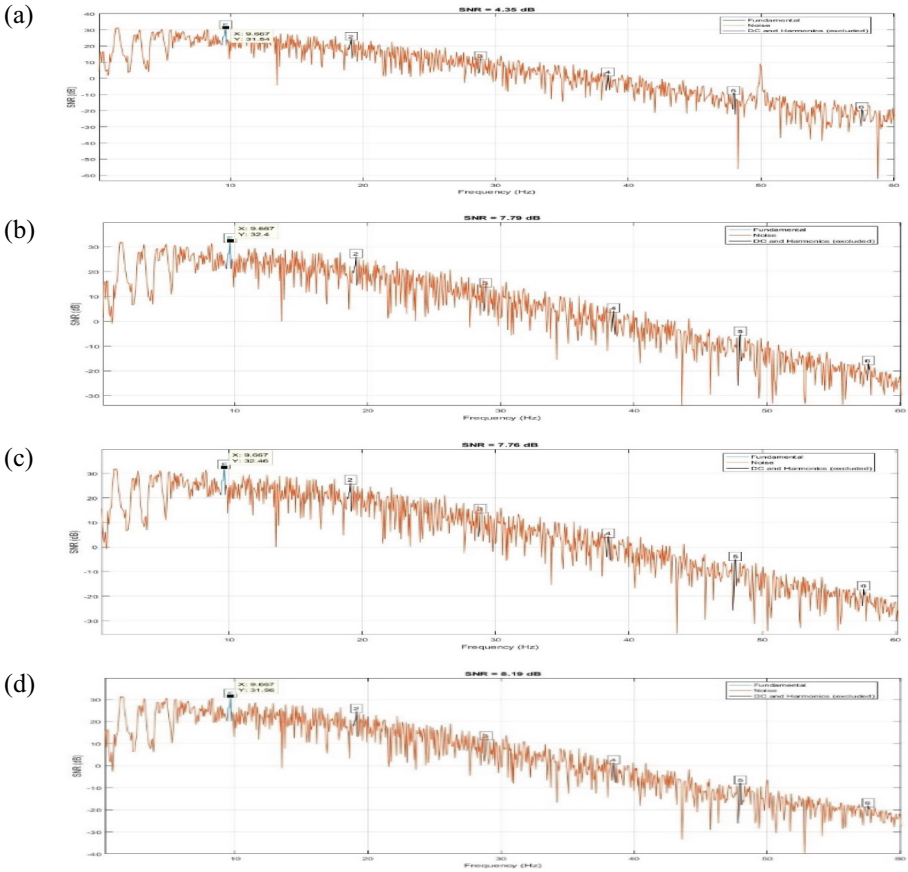


Fig. 4. The comparison of SNR graph digital filter method with Fig. 4(a) SNR graph of raw signal Fig. 4(b) SNR graph of Kaiser window method Fig. 4(c) SNR graph of rectangular window method Fig. 4(d) SNR graph of Tukey window method.

Figure 4(a) shows the SNR graph of the raw ECG signal before filtering. The blue graph represents the fundamental signal, which is the original signal component intended

for detection or measurement. This fundamental signal has a frequency of 9.6 Hz. The red graph illustrates the noise signal, which dominates at a frequency of 50 Hz in the figure. Figure 4(b) shows the SNR graph of the filtered signal using the Kaiser window. The signal shown in blue is the fundamental signal, which refers to the original signal component that you want to detect or measure among all the signals. The fundamental signal is at a frequency of 9.6 Hz, while the noise signal is shown with a red graph. The noise signal at a frequency of 50 Hz can be suppressed so that the noise at that frequency does not become dominant. Figure 4(c) shows the SNR graph of the filtered signal using a rectangular window. The signal shown in the blue graph is the fundamental signal, which refers to the original signal component to be detected or measured among all signals. The fundamental signal is at a frequency of 9.6 Hz. While the noise signal is shown with a red colored graph. The noise signal at a frequency of 50 Hz can be suppressed so that the noise at that frequency does not become dominant. Figure 4(d) shows the SNR graph of the filtered signal utilizing the Tukey window. The signal shown in the blue graph is the fundamental signal, which refers to the original signal component to be detected or measured among all signals. The fundamental signal is at a frequency of 9.6 Hz. While the noise signal is shown with a red colored graph. The noise signal at a frequency of 50 Hz can be suppressed so that the noise at that frequency does not become dominant.

Then, to determine the best digital filter method in reducing noise, a comparative analysis of the SNR value of each filter method is carried out. The data is entered into the Matlab application, then the SNR value of each filter method is calculated, and then each filter is tested for several orders, namely 10, 20, 30, 40, and 50. A high SNR value indicates that the desired signal is significantly stronger than the noise. It indicates that the signal quality is good and the information received is clearer. Whereas a low SNR value indicates that the desired signal is almost comparable to or even weaker than the noise signal. Signaling that the signal quality is poor, and less information is obtained. Then the SNR calculation results of each filter method are shown in Table 1.

Table 1. Comparison of SNR value in the Kaiser, rectangular, and Tukey window methods, each tested at several order levels.

Raw EKG (SNR = 4.35 dB)	Kaiser (dB)	Rectangular (dB)	Tukey (dB)
Orde 10	8.49	7.76	8.19
Orde 20	6.26	6.24	6.61
Orde 30	6.31	6.11	6.48
Orde 40	6.41	6.40	6.44
Orde 50	6.92	6.91	7.01

Table 1 shows the comparison of SNR values of three digital filter methods, namely the window Kaiser, rectangular window, and Tukey window. The results of the SNR value are compared with each digital filter method, and each method uses orders 10, 20, 30, 40, and 50. The best filter method with the highest SNR value is the 10^{th}-order window Kaiser method with a value of 8.49 dB.

Beat per Minute (BPM) is the measurement of heart rate or heart rate in one minute, which is displayed as an electrical signal on an ECG chart. In the ECG signal graph, BPM means the number of R signals per minute. Calculating this BPM value can help in the diagnosis of heart signals. The module is then tested with a phantom ECG comparison tool. The module will be given input from the phantom ECG by providing signals and Beat Per Minute (BPM) values. Then the signal and BPM value will appear on the module monitor screen. In Fig. 5, the module is tested using an ECG phantom.

Fig. 5. Testing the module with the ECG phantom

The module is connected to an ECG phantom using ECG electrode wires, and the ECG signal is shown in Fig. 5. The module will be tested using ECG signals from the phantom using leads II (RA, LL, and RL) with different BPM settings. The results of setting the BPM value from the phantom ECG will be compared with the results that come out on the module monitor screen. Each BPM value setting is recorded six times, and then the results will be averaged. The results of the BPM value test will be displayed in Table 2.

Table 2 presents the comparison results of BPM values between the ECG phantom and the module. Each BPM value setting was tested six times and averaged. The highest error percentage was observed for the BPM setting of 15, yielding a percentage error of 2.222%. Conversely, the lowest error percentage was recorded for the BPM setting of 160, with a percentage error of 0.3125%. The highest distribution value was obtained when setting the BPM value of 120 beats per second. From the results of the comparison of BPM values above, it is concluded that the module still meets the standard of ± 5 BPM. The average obtained at each BPM setting is no more than ± 2 BPM. Then the module will be compared with the patient monitor tool to get the difference in heart rate values. Data was collected 15 times with respondents sitting still for 5 min. The comparison results are listed in the table below. Then, for the t-test results, the p-value obtained is greater than the alpha significance level (0.05), indicating that there is no significant difference between the two groups tested. Table 3 presents the T-Statistic which is lower than the Critical T value and also indicates the statistical insignificance of the difference. From Table 3, it is explained that the results of statistical calculations obtained a p value = 0,43, which means that $p > 0.05$.

Table 2. HR Measurement results in the subject.

Setting ECG Phantom (Bpm)	Module average score (Bpm)	Error (Bpm)	Standard deviation (STD)
15	14.67	2.222	0.516
20	19.83	0.833	0.408
25	24.5	2	0.5477
30	29.67	1.111	0.516
40	39.83	0.416	0.408
60	59.33	1.111	0.516
80	79.17	1.041	0.752
100	99.17	0.833	0.752
120	119	0.833	0.8944
140	139.5	0.357	0.836

Table 3. t-test equal variance

Alpha (0.05)	Modul	Patient monitor
Mean	78.87	78.73
Variance	3.70	4.35
Observations	15.00	15.00
Pooled variance	4.02	
Hypothesized mean difference	0.00	
df	28.00	
t Stat	0.18	
P(T < =t) one-tail	0.43	
t Critical one-tail	1.70	
P(T < =t) two-tail	0.86	
t Critical two-tail	2.05	

4 Discussion

This research uses an ECG circuit consisting of a basic instrument, a BPF circuit, a Notch Filter, an amplifier, and an adder. The body signals that have been tapped using the circuit are connected to the Arduino for data reading. From the results of data reading, signal processing is carried out using the MATLAB application for real-time filtering. Offline analysis using MATLAB aims to find the best filter with the highest SNR value. The methods used are Kaiser, rectangular, and Tukey digital FIR window filters. Of the

three filters, the SNR value is analyzed in various orders. By using order 10, the window Kaiser method excels with an SNR value of 8.49 dB. At order 20, the Tukey window reaches the highest value with 6.61 dB. Then at order 30, window Tukey achieved the highest value with an SNR value of 6.48 dB. In the 40^{th}-order comparison, the Tukey window excels with a value of 6.44 dB. For the 50^{th}-order comparison, the filter with the Tukey method excels with a value of 7.01 dB. From the comparison of filter methods and orders, the highest SNR value is obtained using the 10^{th}-order window Kaiser method with a value of 8.49 dB. Then the best filter is embedded into the system to reduce the interference signal. Then the tool is tested on respondents to see the results of the filter performance. It is proven that the signal displayed is clean from noise interference. Subsequently, the system underwent testing using an ECG signal input, adjusting the BPM value using a phantom ECG comparison tool. Each BPM setting was repeated six times to calculate the average value, percentage error, and standard deviation. The highest error percentage was observed for the BPM setting of 15, at 2.222%. Conversely, the lowest error percentage was found for the BPM setting of 160, with an error rate of 0.3125%. From the comparison of BPM values, it is concluded that the module still meets the standard of ± 5 BPM. The average obtained at each BPM setting is no more than ± 2 BPM. Then, for the t-test results, the p-value obtained is greater than the alpha significance level (0.05), indicating that there is no significant difference between the two groups tested. The T-statistic, which is lower than the Critical T value, also indicates the statistical insignificance of the difference.

In a study conducted by Chhavi Saxena et al. in 2018, the performance of FIR and IIR filters was analyzed. The research indicated that FIR filters significantly suppress high-frequency components more effectively compared to IIR filters. The standard deviation of the IIR filter output closely matches the desired output, whereas the FIR filter output exhibits lower values in the presence of high-frequency noise compared to the IIR filter. Overall, the FIR filter demonstrated more accurate results for the proposed algorithm compared to the IIR filter [13]. Additionally, research conducted by Amir Rasel et al. in 2021 focused on the design of window-based FIR filters for noise reduction in ECG signals. The study explored FIR filter designs using various windows such as Rectangular, Kaiser, Hamming, and Bartlett. Then the SNR value of these methods is calculated and compared. It was found that Kaiser Window has the highest SNR value of the other methods, with an SNR value of 0.0064 [14]. Further research by Bhaskar et al. in 2016 conducted research on high-frequency electromyograph signal noise removal using an FIR low-pass filter. The approach employed involved utilizing a FIR digital filter based on Distributed Arithmetic (DA) architecture. The study calculated the SNR of the processed ECG signal and determined the Mean Squared Error (MSE) to assess the discrepancy between the noise and the filtered ECG. A low-pass FIR filter designed with the Kaiser Window proved highly effective in attenuating high-frequency EMG interference from noisy ECG signals using a low-order filter. The SNR achieved with the Kaiser Window was 0.6438 dB, and the MSE was 0.0501. These results outperformed those obtained with the Rectangular, Hanning, Hamming, Blackman, and Bartlett Window methods [15]. Additionally, in a study by K. Sravan Kumar et al. in 2015, research focused on ECG noise reduction using FIR and IIR digital filters employing various methodologies. The FIR filter utilized windowing techniques such as Rectangular, Hann, Hamming,

Blackman, and Kaiser. Conversely, the IIR filter employed types including Butterworth, Chebyshev I, Chebyshev II, and Elliptic. Evaluation of the filters was based on waveform analysis, power spectrum density (PSD), signal-to-noise ratio (SNR), and Mean Squared Error (MSE). Among the FIR filters, the Kaiser window with an order of 56 demonstrated superior performance compared to other orders in both FIR and IIR filters at orders 1, 2, and 3. Specifically, the Kaiser window method exhibited better SNR values compared to alternative FIR filter methods. Then the comparison in terms of MSE showed that the average Kaiser window method obtained a lower MSE value than other methods for each amount of data [16].

The lead II ECG signal collected from the body is subsequently analyzed using MATLAB to identify the most effective digital filter for noise reduction in the ECG signal. From the analysis of the comparison of signal shape and SNR value, a filter with the best SNR value is obtained in the Kaiser window method using order 10. The digital filter is embedded in the Arduino microcontroller, then the data will be sent to the Raspberry using serial communication to display the signal on the monitor screen. And the signal displayed can appear in real time without any slight delay. Then, the test was conducted using the respondent's body, and the results of a clean ECG signal were obtained, but when the respondent moved, the signal displayed noise. The weakness of this research digital filter has not been able to reduce the signal due to body movement, so when experiencing movement, the heart signal that appears on the monitor screen is disturbed.

This research aims to develop a bedside monitor equipped with a digital filter specifically designed to reduce noise in ECG signals. With this development, it is expected that the bedside monitor can produce a cleaner ECG signal, improve monitoring accuracy, and ultimately, provide significant benefits in patient care. Significant benefits in patient care. In addition, this research will also evaluate the effectiveness of the digital filter used, both in terms of the quality of the generated signal as well as from the perspective of clinical use in the field.

5 Conclusions

In this study, we developed a bedside monitor that integrates a digital filter to reduce noise in ECG signals. The second-lead ECG signal obtained from the body is analyzed using MATLAB to determine the optimal digital filter to minimize noise in the ECG signal. From the analysis of the comparison of signal shape and SNR value, a filter with the best SNR value is obtained in the Kaiser window method of order 10. The performance of the filter can outperform the performance of filters with other types. These results are seen from the Fast Fourier Transform (FFT) process, SNR value, and the result of the signal. The digital filter is embedded in the Arduino microcontroller, then the data will be sent to the Raspberry using serial communication to display the signal on the monitor screen. This tool is tested using a comparison in the form of an ECG phantom, then given an ECG signal with additional noise with a frequency of 50 Hz. The signal displayed on the monitor screen looks clean without the PLN noise signal. Then tested using the respondent's body, the results of the ECG signal were clean, but when the respondent moved the signal displayed experienced noise. For future research, it is recommended

to explore the ECG signal filter to reduce the movement signal. Other filter methods can be used to reduce the movement signal.

References

1. Prastyadi C, Irianto BG, Ariswati HG, Titisari D, Nyatte S, Misra S (2022) Analysis of the accuracy of temperature sensors at the calibrator incubator laboratory are equipped with data storage base on Internet of Thing. Indones J Electron Electromed Eng Med Informatics 4:160–167. https://doi.org/10.35882/ijeeemi.v4i3.241
2. 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 14:811. https://doi.org/10.14716/ijtech.v14i4.5678
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 Informatics 3:53–58. https://doi.org/10.35882/ijeeemi.v3i2.3
4. Rachman RA, Wisana IDGH, Nugraha PC (2021) Development of a low-cost and Effisient ECG devices with IIR digital filter design. Indones J Electron Electromed Eng Med Informatics 3:21–28. https://doi.org/10.35882/ijeeemi.v3i1.4
5. Francis J (2016) ECG monitoring leads and special leads. Indian Pacing Electrophysiol J 16:92–95. https://doi.org/10.1016/j.ipej.2016.07.003
6. Kumar S, Mehra R, Chandni (2016) Implementation and designing of filters using Kaiser window for denoising of electrocardiogram signals on FPGA. In: IEEE International conference on computer communications control (IC4 2016), p 128
7. Velayudhan A, Peter S (2016) Noise analysis and different denoising techniques of ECG signal—a survey. IOSR J Electron Commun Eng 3:40–44
8. Patwary AB, Chowdhury MTI, Mamun N (2014) Comparison among ECG filtering methods for nonlinear noise. In: 2018 International conference on advancement in electrical and electronics engineering, vol 4, pp 23–34
9. Usharani BJRA (2015) Enhancement of ECG signal by using digital FIR filter. Int J Sci Res 4:910–913
10. Mittal A, Rege A (2016) Design of digital FIR filter implemented with window techniques for reduction of power line interference from ECG signal. In: IEEE International conference on computer communications control IC4 2015, pp 2–5. https://doi.org/10.1109/IC4.2015.7375535
11. Buendía-Fuentes F et al (2012) High-bandpass filters in electrocardiography: source of error in the interpretation of the ST segment. ISRN Cardiol 2012:1–10. https://doi.org/10.5402/2012/706217
12. Blanco-Velasco M, Weng B, Barner KE (2008) ECG signal denoising and baseline wander correction based on the empirical mode decomposition. Comput Biol Med 38:1–13. https://doi.org/10.1016/j.compbiomed.2007.06.003
13. Saxena C, Sharma A, Srivastav R, Gupta HK (2018) Denoising of ECG signals using FIR & IIR filter: a performance analysis. Int J Eng Technol 7:1–5. https://doi.org/10.14419/ijet.v7i4.12.20982
14. Ameer Rasel VB, Uma A, Kalyani P (2022) Design of window based FIR filter for electrocardiogram noise removal. Int J Nonlinear Anal Appl 12:1519–1528. https://doi.org/10.4108/eai.7-12-2021.2314495
15. Bhaskar PC, Uplane MD (2016) High frequency electromyogram noise removal from electrocardiogram using FIR low pass filter based on FPGA. Procedia Technol 25:497–504. https://doi.org/10.1016/j.protcy.2016.08.137

16. Kumar M (2024) Design and implementation of digital low pass FIR and IIR filters using VHDL for ECG denoising. Int J Eng Trends Technol 72:252–265. https://doi.org/10.14445/22315381/IJETT-V72I1P125
17. Gawande GS, Khanchandani M (2012) Performance analysis of FIR digital filter design techniques. Int J Comput Corp Res 2
18. Manlong NA, Rahul J, Sora M (2018) ST segment analysis for early detection of myocardial infarction. Int J Comput Sci Eng 6:1500–1504. https://doi.org/10.26438/ijcse/v6i6.15001504
19. Bui NT et al (2020) Real-time filtering and ECG signal processing based on dual-core digital signal controller system. IEEE Sens J 20:6492–6503. https://doi.org/10.1109/JSEN.2020.2975006
20. Luo S, Johnston P (2010) A review of electrocardiogram filtering. J Electrocardiol 43:486–496. https://doi.org/10.1016/j.jelectrocard.2010.07.007
21. Radhika B, Sheshagiri Rao VR (2019) Incubator baby parameter sensing and monitoring. Int J Innov Technol Explor Eng 8:2945–2947
22. Prayogo I, Alfita R, Wibisono KA (2017) Monitoring system for heart rate and body temperature as an IoT (Internet of Thing)-based patient health level indicator using the fuzzy logic method using android. Triac Electr Comput Eng J 4
23. Obeidat YM, Alqudah AM (2023) An embedded system based on raspberry pi for effective electrocardiogram monitoring. Appl Sci 13. https://doi.org/10.3390/app13148273
24. Reddy VV, Neelaveni M, Sandeep M, Ahmad SK, Krishna PV (2023) Comparison of fir and Iir filters using ECG signal with different sampling frequencies. Int Res J Mod Eng Technol Sci 05:2576–2581. https://doi.org/10.56726/irjmets35977
25. Hamici H, Kanan A, Al-hammuri K (2023) Optimized FIR filter using genetic algorithms: a case study of ECG signals filter optimization. BioMedInformatics 3:1197–1215. https://doi.org/10.3390/biomedinformatics3040071
26. Chavan MS, Agarwala R, Uplane MD (2008) Design and implementation of digital FIR equiripple notch filter on ECG signal for removal of power line interference. WSEAS Trans Signal Process 4:221–230
27. Saritha Raj K, Rajesh Kumar P, Satyanarayana M (2023) Capricious digital filter design and implementation using Baugh–Wooley multiplier and error reduced carry prediction approximate adder for ECG noise removal application. Circ Syst Signal Process 42:6726–6748. https://doi.org/10.1007/s00034-023-02417-z
28. Mir HY, Singh O (2024) Powerline interference reduction in ECG signals using variable notch filter designed via variational mode decomposition. Analog Integr Circ Signal Process 118:317–328. https://doi.org/10.1007/s10470-023-02200-9
29. Bhateja V, Verma R, Mehrotra R, Urooj S (2013) A non-linear approach to ECG signal processing using morphological filters. Int J Meas Technol Instrum Eng 3:46–59. https://doi.org/10.4018/ijmtie.2013070104
30. Srinivasa MG, Pandian PS (2019) Elimination of power line interferance in ECG using adaptive filter, notch filter, and discrete wavelet transform techniques. Int J Biomed Clin Eng 8:32–56
31. Jain S, Paul S (2022) Design of filters using current amplifiers for removal of noises from ECG signal. Procedia Comput Sci 218:1888–1904. https://doi.org/10.1016/j.procs.2023.01.166
32. Azzouz A et al (2024) An efficient ECG signals denoising technique based on the combination of particle swarm optimisation and wavelet transform. Heliyon 10:e26171. https://doi.org/10.1016/j.heliyon.2024.e26171
33. Thannoon HH, Hashim IA (2023) Efficient enhanced recursive Least Square algorithm adaptive filtering scheme for artifacts removal in ECG signals. e-Prime—Adv Electr Eng Electron Energy 6:100–318. https://doi.org/10.1016/j.prime.2023.100318

34. Munoz-Macho AA, Dominguez-Morales MJ, Sevillano-Ramos JL (2024) Analyzing ECG signals in professional football players using machine learning techniques. Heliyon 10:e26789. https://doi.org/10.1016/j.heliyon.2024.e26789
35. Tabakov S, Iliev I, Krasteva V (2008) Online digital filter and QRS detector applicable in low resource ECG monitoring systems. Ann Biomed Eng 36:1805–1815. https://doi.org/10.1007/s10439-008-9553-5

Tool Wear Fault Detection and Classification Using Vibration Signal Analysis

Raja Siti Nur Adiimah Raja Aris[1,2(✉)], Fahmi Samsuri[1], Siti Nurshafina Zaharah Saffinee[2], Damhuji Rifai[2], Kharudin Ali[2], and Wasis Nugroho[2]

[1] Faculty of Engineering Technology, University Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia
nuradiimah@uctati.edu.my
[2] Faculty of Engineering Technology, University College TATI, 24000 Teluk Kalong, Kemaman, Terengganu, Malaysia

Abstract. This paper presents a real-time tool wear performance monitoring system from data collection and data analysis to real-time tool wear fault classification. The lack of a monitoring system approach in tool wear fault detection for CNC machine applications leads to the effect on tool lifetime and production effectiveness. To overcome this issue, the tool wear fault is approached using a vibration sensor embedded in the spindle of the CNC machine. The state of the cutting edge is assessed based on wear observed on the flank surface during the milling processes. The experiment method in this research employed High-speed steel (HSS) end mills as the cutting tools. The tests were conducted with two different tools under combinations of the cutting speed (800 m/min) and the depth of cut varies at 0.5 mm and 1 mm. The vibrations due to the flank wear were measured using piezoelectric sensors embedded within a spindle of a CNC machine and the vibration data was recorded using a reliable data acquisition (DAQ) device during the milling process. Additionally, the tool wear image is captured using the microscope and the flank surface of tool wear is measured to analyze the tool wear lifetime. The method for signal analysis is carried on with the implementation of the time domain analysis and frequency domain analysis, Fast Fourier Transform (FFT) representation for vibration signal. Based on the result achieved, the amplitude of the vibrations increased with increasing the depth of cut and the flank wear value. The lifetime of the tool achieved accelerated wear during cycle 6 for Tool 1 (0.6425 mm) and Tool 2 (0.9615 mm) of the experiment.

Keywords: CNC machine tool · Fault diagnosis · Vibration signal

1 Introduction

Maintaining and enhancing machining procedures is essential in today's industrial environment to guarantee successful and efficient operations. The accurate monitoring and promptly identifying the cause of tool wear during machining processes is one of the

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 336–360, 2026.
https://doi.org/10.1007/978-981-95-9233-3_25

biggest issues faced by manufacturers. Over the past few years, researchers have extensively investigated tool wear monitoring technology. The detection principle can be categorized into two distinct methods: indirect and direct [1, 2]. The indirect method involves estimating the tool wear state by monitoring sensor signals that reflect the tool wear status. These signals include cutting force, acoustic emission, spindle motor current and power, workpiece surface texture, and tool/workpiece vibration [3–5]. Vibration signals are also one of the most important parameters for the measurement of tool wear by using the indirect sensing method. Machining with worn tools increases the friction between the tool and the workpiece because of these internal fractures of the tool takes place. Due to these force fluctuations, vibration occurs in the system [6]. The condition of the cutting tool during machining stands out as a crucial parameter, greatly influencing the tool's effectiveness in the machining process. The significance of tool wear lies in its impact on both the tool's life and the quality of the machined hole [7, 8].

From the observation [9] and [10], the lack of a monitoring system approach in tool wear fault detection for CNC machine application leads to the effect on production effectiveness. The traditional way of tool wear detection is normally through inspection by experienced technical personnel. This technique involves human subjective intervention in qualitative analysis which tends to lead to the low quality of the analysis's performance [11]. Manual judgment practices also are the main cause of the breakdown of CNC machine devices especially in the conventional manufacturing industry [12]. To overcome this issue, excessive tool wear monitoring is essential in the manufacturing sector. Besides, there are a few implications from the measurement issue such as it will reduce the precision and quality of the machine parts, increased cost involving tool replacements, production lines reschedules, and extra safety precaution arrangements while the machine is being checked.

Piezoelectric sensors are frequently utilized in industry research and development because they are a flexible instrument for measuring signals during machining operations [13]. Similarly, vibration sensors are among the most commonly used measurements for detecting tool wear [14, 15]. Vibration sensor feature monitoring is cost-effective, easy to set up, and offers a periodic pattern similar to the cutting force, making it widely used in tool wear monitoring [16]. The advanced features of vibration signals proved to be highly effective in detecting drill tool failure and wear while remaining unaffected by variations in cutting conditions [17]. The use of a piezoelectric vibration sensor that offers a sophisticated transducer that converts mechanical vibrations into electrical signals is essential to this endeavor. Due to its great sensitivity and ability to acquire data in real time, the sensor provides a special chance to record the minute details of vibrations produced during milling operations.

The researcher [18–20] used the vibration signal to monitor the cutting tools during the machining operation. The result from the output sensor will be analyzed on the relationship between the tool wear and vibration values. Once the process parameter has been sensed, it is essential to employ an appropriate signal-processing technique to extract important information regarding the tool condition. The use of bidirectional LSTM models for cutting tool wear estimation shows high accuracy, optimizing technological processing and reducing downtime when integrated into machine control systems. Additionally, the deep learning approaches eliminate manual feature selection, making

them suitable for large datasets, while extending tool life, improving cut quality, and ensuring timely replacement before failure. High classification accuracy, often above 80%, is achieved, particularly in detecting critical wear transitions. In fault detection, methods that incorporate more sensor signals show promise in enhancing reliability and identifying machine tool spindle conditions.

However, the challenges remain such as the need for well-representative training data to avoid misclassification. Misclassified cases, while often labeled as failures leading to tool replacement, may not be optimal for efficiency. There is also the issue of unknown faults being misclassified as known ones if they haven't been learned, and the difficulty of transferring these models between different machines, even those of the same type. Additionally, some methods consider only limited fault conditions, which could restrict their applicability.

In [21] research, the vibration signal is applied to diagnose tool wear during end milling of difficult-to-cut AMC with 10% SiC content. The measurement of vibration signal has been implemented using a Triaxial piezoelectric charge accelerometer Type 4321 Brüel and Kjær with three different directions and operated temperature range between 250 °C and measure up to 10,000 Hz during the experiment. Various signal processing methods are available, each being highly subjective and specifically tailored to the application. This approach is effective when traditional regression models fall short and is particularly relevant to the automotive and aerospace industries, where AMCs reinforced with SiC or Al_2O_3 are commonly used. The research employs a diamond-coated three-edged end mill for milling tests, incorporating vibration acceleration and cutting forces as input data across three different cutting speeds, and assesses model efficiency using mean square error (MSE). However, reliance on ANN models highlights the limitations of simpler predictive techniques, and the complex model selection process may introduce variability in results. The focus on AMC materials with 10% SiC content may restrict the generalizability of the findings to other compositions or cutting conditions. Additionally, depending solely on MSE as an evaluation metric may not fully capture the models' overall performance and accuracy.

Regarding [22], the system with a cost-effective that offers a multi-sensor application for tool conditioning monitoring is suitable for SMEs' applications to overcome the cost of operation and limited system of monitoring. In this research, three types of sensors are implemented to monitor cutting force, energy consumption, and vibration signals during the milling process. Data will be collected and analyzed using big data technology through Apache Spark. The Fast Fourier Transform (FFT) is the most commonly used signal analysis method [23]. The researcher [24] studied vibration analysis under three different conditions during the milling operation a new tool medium-wear tool, and a blunt tool. All the vibration results were recorded as frequency and acceleration using an FFT analyzer. Besides, other researchers also apply an FFT analyzer to measure the vibration of the cutting tool during machine operations [6]. In this research, the vibration sensor is implemented in measuring tool wear faults. A vibration sensor is used as a fault indication variable based on its capability to enhance the accuracy and reliability of measurements. Additionally, the research will analyze the tool wear image by using the digital microscope to measure the tool wear breakage level. The data collected from the vibration sensor will analyzed using the time domain and frequency domain analysis.

2 Materials and Methodology

2.1 Piezoelectric Vibration Sensor

The piezoelectric vibration sensor in Fig. 1 is a type of transducer that converts mechanical vibrations or accelerations into electrical signals. The piezoelectric vibration sensor is an essential component in vibration monitoring for tool wear on a CNC machine because it captures and transforms vibrational energy generated during milling into electrical signals. A voltage signal is then produced by measuring this electric charge. Due to its vibration sensitivity, the sensor can identify even minute changes in mechanical conditions, which makes it useful for monitoring tool wear during milling operations. In this experiment, the sensor is attached to the top side of the CNC machine's tool holder.

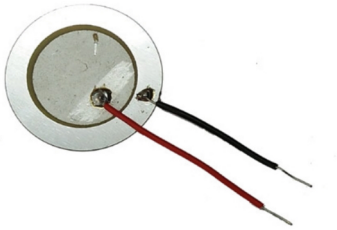


Fig. 1. Piezoelectric Vibration Sensor

2.2 MCC USB-128FS-PLUS Data Acquisition (DAQ) Device

A Data Acquisition (DAQ) device as in Fig. 2, is a specialized hardware component designed to gather, measure, and record analog or digital data from various sensors and transducers. Within the framework of the project concerned with vibration monitoring for tool wear on a CNC machine, the electrical signals produced by the piezoelectric vibration sensor are captured and processed using a DAQ device. USB-128FS-PLUS features include a 50kS/s sample rate, eight analog inputs, two (2) 12-bit analog outputs, 16 digital I/O connections, and one (1) 32-bit external counter. To monitor this device's data, DAQami software is used. DAQami is optimized for data logging, visualization, and signal generation. Data is viewed in real-time or post-acquisition. The device used in this project has a particular channel designated specifically for taking in data from the piezoelectric vibration sensor.

Fig. 2. Data acquisition (DAQ) device

2.3 RIEVBCAU MS3 Digital Microscope

A digital microscope is an advanced optical tool as in Fig. 3 is used to record enlarged images and videos and display them on a linked computer, in contrast to traditional microscopes that rely on eyepieces for observation. RIEVBCAU MS3 Digital Microscope with 1500X magnifier and nine (9) Inch display screen is utilized in this experiment. The screen can be rotated at 90-degree angles for different needs and is more practical. The digital microscope offers 12 Megapixels which gives ultra-precise focusing and a wired remote controlled for a more stable and wider view. Tool wear image measuring and capturing are applied with HiView software. HiView offers real-time viewing without delay and HD resolution. The tool is being put under the microscope and the display is either directly on the microscope LCD or through HiView software to observe the tools.

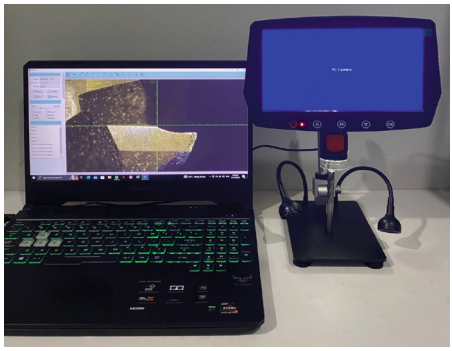


Fig. 3. RIEVBCAU MS3 Digital Microscope

2.4 The Proposed Fault Detection System for Tool Wear

In this study, the vibration sensor is used to measure the fault of the tool wear. Figure 4 shows the architecture design fault detection system for tool wear. The design is divided into four sections. The first section is the power supply. The AC supply will be converted to the DC supply and used for circuit controllers. The second section will discuss excitation signal generators. The excitation signal is used as an input signal for the piezoelectric vibration sensor.

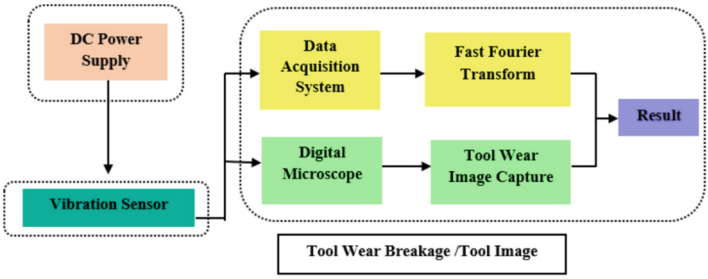


Fig. 4. Fault Detection System for Tool Wear

The next section is the method to control or process the devices. This section will implement the Data Acquisition System (DAQ) as a data-receiving and filtering device. The DAQ is connected to a PC for transferring data collection from a sensor signal. The last section present was tool wear fault measurement. This method will apply PC usage to process the data from the DAQ devices and plot the graph based on the data computed. Besides, the tool wear image will be captured using the digital microscope to measure the tool wear breakage level. The fault detection system for tool wear design acted as the main system in this research. The signal generator will activate a sensor to produce a vibration signal. For DAQ activation, the 5VDC is used by connecting the DAQ directly to the PC through a USB port for data transfer. The analog pin is used for the sensor connection on positive polarity meanwhile the negative polarity will connect to the ground pin.

The data analysis process begins when the sensors are used for measurement of fault detection on the tool wear and the sensor signal will be sent to the PC through the DAQ as a data interface. The software that is used to display and analyze the signal is DAQami software. The DAQami software presents the raw data through graphs and values, allowing for both real-time observation and recording for subsequent offline analysis. The data collected will be compared with the conventional meter for the validation process. Once this process is achieved, the classification of tool wear will be examined using the Fast Fourier Transform (FFT) method. The output signal from the sensor shows a set of data for tool wear classification. Besides, the tool wear image also will be captured using the digital microscope to measure the level of tool damage. The measurement of tool wear life will be referred to the valid reference, the International Standards Organization (ISO-8688-1/1994 standard).

2.5 Experiment Setup

The experiment will be conducted on a three-axis CNC turn-mill machine tool (NL25000SY/700), as shown in Fig. 5. The experiment is conducted with 9 cycles of each set for Tool 1 and Tool 2 to collect vibration data with various parameters; speed (800 rpm), depth of cut (0.5 mm and 1.0 mm) and fixed rate. The workpiece material mild steel in the category of low Carbon Steel is selected, as shown in Fig. 6. The selected cutting tool is an HNT Solid Carbide Long Endmill 4 Flutes tool and no coating, as shown in Fig. 7.

Fig. 5. Mori Seiki (NL2500SY/700) CNC Machine

Fig. 6. Workpiece dimension

Fig. 7. HNT Solid Carbide Long Endmill 4 Flutes tool

In this study, vibration sensors are used, and the mounting locations are shown in Fig. 8. The vibration signals will be collected from a piezoelectric vibration sensor that is mounted on the spindle. The process of handling feedback data in an experiment involving the use of a Data Acquisition (DAQ) module, as shown in Fig. 9, specifically the USB 128FS-PLUS model.

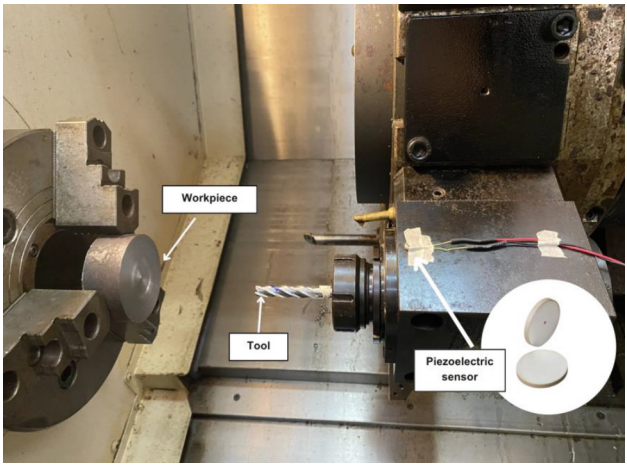


Fig. 8. The piezoelectric position on the spindle

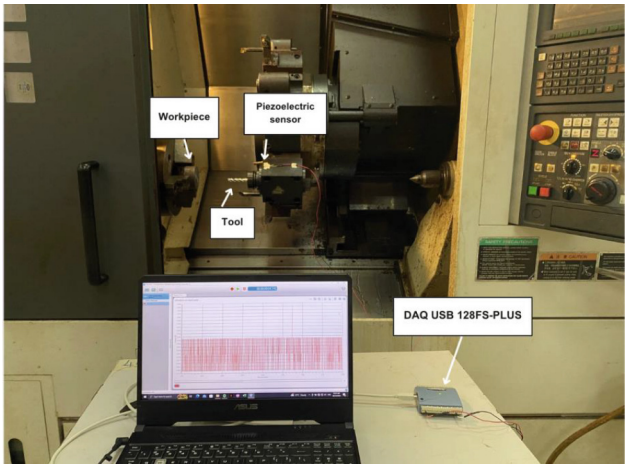


Fig. 9. The Experimental Setup

The DAQ module is equipped with both Analog Input/Output (I/O) and Digital I/O capabilities, making it versatile for processing various types of signals. Then, the feedback data signals are collected and processed by the DAQ module. The processed

data is then transmitted to a personal computer (PC) through the USB port. This step involves transferring the data from the experimental setup to a more accessible and user-friendly computing environment for further analysis. The processed data on the PC is compared with data from a meter. The meter serves as a reference or standard against which the DAQ-processed data is checked, allowing for the validation and verification of the experimental results. Next, the time domain and fast Fourier transfer method will be employed to examine the data and make classifications related to the condition of tools, particularly to identify instances of tool wear. This type of analysis is common in various industries, especially those involving manufacturing and machining, where monitoring and predicting tool wear are crucial for maintaining efficiency and minimizing downtime. Meanwhile, the tool wear analysis is examined for analysis of tool wear lifetime. Tool wear measurement setup using an MS3 Screen digital microscope, and a ring light will be installed within the machine tool, depicted in Fig. 10. The HiView image processing software will be employed to determine the distance between the unworn contour and the maximum wear point.

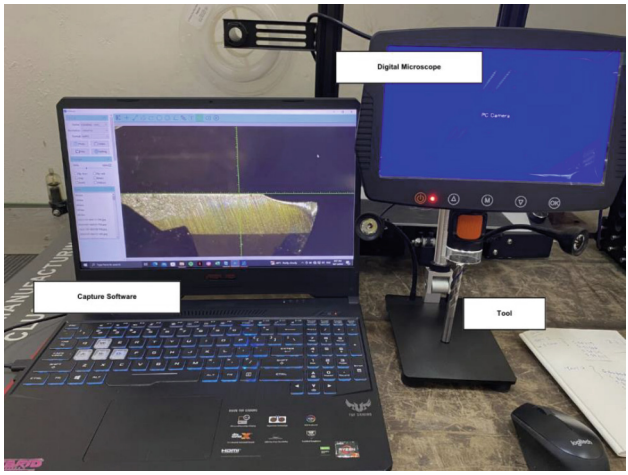


Fig. 10. Microscope Setup

2.6 Tool Wear Measurement Setup

Tool wear measurement is an essential aspect of maintaining optimal performance and efficiency in machining processes, and various methods are employed to assess and monitor the condition of cutting tools. Visual inspection is a direct approach that examines the tool for observable signs of wear, such as flank wear or edge rounding. For more precise measurements, toolmaker microscopes and Coordinating Measuring Machines (CMM) can be employed to analyze tool geometry more accurately. For this research, the MS3 Screen digital microscope is used to measure tool wear after the machining process at time intervals, as shown in Fig. 11. As shown in Fig. 12, the uniform flank wear (VB1) is measured over the tool's whole cutting edge. The average of these measurements is

recorded as the tool's final wear. In some cases, repeating the measuring procedure several times helps to precisely determine the degree of uncertainty in the measurements. Initially, the tool is removed from the machining equipment and cleaned to eliminate any residues or contaminants. The next step is mounting the tool on a suitable device, such as a microscope slide, ensuring stability and optimal accessibility for examination. Setting up the microscope involves selecting an appropriate magnification level, adjusting lighting conditions, and calibrating as needed. Once the microscope is ready, detailed images or videos of the tool surface are captured using a digital camera attached to the microscope. Multiple angles are often explored to provide a comprehensive view, and relevant information, including tool type and cutting conditions, is documented.

Subsequently, the captured images undergo analysis using image analysis software (HiView) to quantify wear characteristics such as wear area and depth. The findings are then documented, creating a valuable reference for future analyses or comparisons between different tools. For post-analysis, it is important to preserve the tool for potential further examinations or archiving purposes. Overall, this method provides a comprehensive understanding of tool wear, aiding in predictive maintenance and improving the efficiency of manufacturing processes.

Fig. 11. MS3 Screen microscope used for tool wear measurement

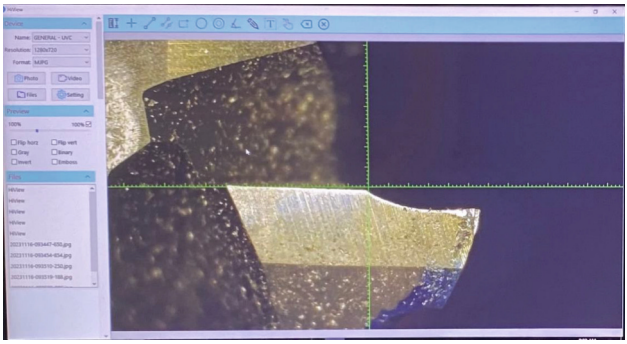


Fig. 12. HiView Software Interface

As a general practice, based on [25–27], flank wear (VB) stands out as the main factor employed to evaluate the lifetime of a cutting tool. The progress of flank wear throughout machining time is disclosed through three stages: the initial wear stage, wear stabilization, and accelerated wear, as visually depicted in Fig. 13. This progression provides a comprehensive understanding of how flank wear develops over time, offering valuable insights into the tool's performance and durability throughout different stages of its operational life. To verify the tool's long life, the International Standards Organization proposes that this investigation rely on the ISO-8688-1/1994 standard for testing and analysis. According to the standard, the tool life is considered achieved when the uniform wear of the flank surface measures 0.3 mm. Additionally, if the surface exhibits peeling, severe grooving, or uneven wear exceeding 0.6 mm, the tool will reach the failure region, signifying the conclusion of its life cycle and necessitating replacement [28–30].

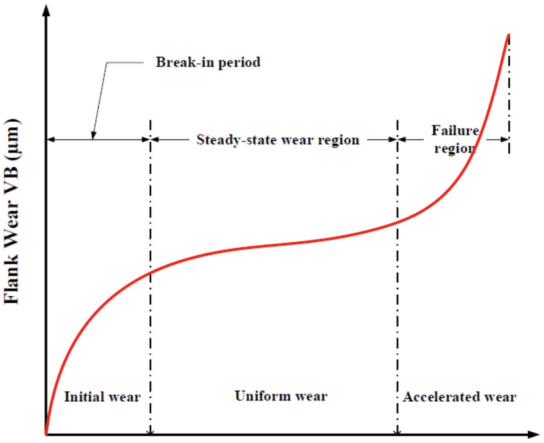


Fig. 13. Reference graph for tool wear lifetime based on the ISO-8688-1/1994 standard

The lifetime of a cutting tool is affected by various uncontrollable factors. During a high-speed operation, the contact surface between the tool and workpiece produces a

significant amount of heat, leading to alterations in the cutting edges through intricate physicochemical phenomena. The lifetime of the tool is defined by the duration it takes to reach.

2.7 Data Analysis Process

Fault classification of tool wear involves the identification and categorization of various types of faults or abnormalities associated with the condition and wear of tools in manufacturing processes. The classification is instrumental in predictive maintenance, quality control, and overall operational efficiency within industrial settings. The types of faults for this research focus on flank wear which consists of initial wear, uniform wear, and accelerated wear. Monitoring techniques encompass a range of sensor data; vibration, as well as tool wear lifetime analysis. For this research, data analysis is vital for fault classification involving time domain analysis, and frequency domain analysis, Fast Fourier Transform (FFT) representation for vibration signal. Time domain analysis refers to the interpretation process of signals or data over time. In tool wear monitoring, signals from vibration sensors are analyzed versus time. The peak amplitude from the vibration signals may indicate tool wear progression. A sudden change in the peak amplitude could signify that the tool is approaching failure.

Meanwhile, the FFT is applied to detect the presence of specific frequencies that correspond to the flank wear of the tools. By transforming a time-domain signal into the frequency domain, it becomes easier to identify periodic signals or detect abnormalities. For example, if the frequency of vibrations increases over time, it can indicate progressive tool wear. Feature extraction from sensor data is crucial for representing the state of tool wear, and pattern recognition is essential for developing models that can identify specific tool wear faults. Meanwhile, the tool wear image will be captured using a microscope involves a systematic process to examine and analyze the condition of a tool's surface, offering crucial insights for manufacturing and materials science. The comparison between tool wear data collection and tool wear image for classification involves a systematic approach that integrates image analysis and data analysis techniques. Initially, numerical tool wear data from sensors, labeled with the corresponding wear states, is collected. Simultaneously, tool wear images are captured using a microscope or another imaging device, covering various wear states and ensuring representation. The next step is data validation where the data collected from the signal will validate with the meter conventional to measure the data.

3 Results

An experimental analysis was conducted to examine the impact of tool wear on vibration signals. A total of 18 experiment cycles were executed to fulfill the objectives of this research. This study employs two methods, namely data acquisition and image inspection, to investigate tool wear. The primary aim of the experiment is to analyze the condition of tool wear. The objective is to verify that when the length difference of tool wear reaches 0.6 mm, the tool wear is visible in both the signal's amplitude and images captured by a microscope. The goal was to establish a correlation between vibration

signals and tool wear, and the results revealed that vibration amplitudes increase with escalating tool wear.

3.1 Tool Wear Measurement Based on Vibration Amplitude

The greatest displacement or distance that a vibrating element or signal can go from its equilibrium position is known as its amplitude, and it is a fundamental metric in vibration monitoring. The strength or intensity of the vibration is measured by it. The maximum amplitude indicates the highest value that the vibrating object can reach while in motion, depending on the situation and the measurement type. Table 1 shows the maximum voltage from each cycle of the experiment. The maximum voltage value is not gradually increasing after every cycle; however, it does show that the last cycle's value is the highest compared to previous cycles for both Tool 1 and Tool 2. Tool 1 resulted in reaching the highest voltage at 0.206 V meanwhile Tool 2 achieved the highest amplitude at 0.274 V. This result can be concluded that the more vibration occurs during the cutting process, the higher the voltage value can be obtained. Vibration patterns can vary because of changes in various characteristics. Higher vibration levels are frequently the outcome of deeper cuts since they frequently result in stronger cutting pressures and more intense interactions between the cutting tool and the workpiece.

Table 1. The amplitude of vibration for Tool 1 and Tool 2

	Tool 1	Tool 2
Cycle:	Voltage(V):	Voltage(V):
1	0.137	0.108
2	0.093	0.074
3	0.093	0.118
4	0.137	0.074
5	0.137	0.015
6	0.015	0.005
7	0.147	0.245
8	0.157	0.274
9	0.206	0.196

3.2 Tool Wear Measurement Using Microscope

The International Standards Organization recommends that a tool should be replaced if its useable life is over, and it has significant grooves on it, or if the uneven wear has reached 0.6 mm or greater. Table 2 shows the measurement of tool wear length for every cycle of the experiment based on Tool 1 and Tool 2. The initial length for Tool 1 is 5.80349 mm and Tool 2 is 6.08983 mm. The length value of each tool was

decreased gradually based on the cutting process during the experiment. The cycle 9 length value for every tool achieved 3.88032 mm and 3.49359 mm as shown in Fig. 14. This figure represents the comparison value of tool wear image between Tool 1 and Tool 2 at different depths of cut varies at 0.5 mm and 1 mm. Increasing the depth of cut value will generate higher cutting force that leads to more significant tool wear and potential tool breakage. It will give an effect on the tool experience greater stress and shorten the tool's lifetime.

Table 2. Tool wear length value for each cycle for Tool 1 and Tool 2

	Tool 1		Tool 2	
Cycle:	Length (mm)	Difference (mm)	Length (mm)	Difference (mm)
Initial	*5.80349*	0	6.08983	0
1	5.73718	0.0663	5.93165	0.1582
2	5.6098	0.1274	5.89744	0.0342
3	5.51366	0.0961	5.80137	0.0961
4	5.48161	0.0320	5.73754	0.0638
5	5.0001	0.2246	5.51291	0.2246
6	4.90395	0.0961	5.12831	0.3846
7	4.77575	0.1282	5.06451	0.0638
8	4.52105	0.2547	4.26294	0.8016
9	3.88032	0.6407	3.49359	0.7694

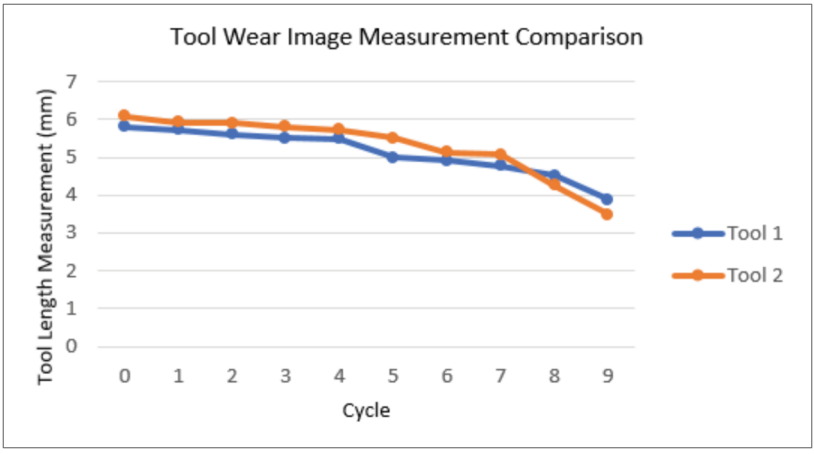


Fig. 14. Tool Wear Image Measurement Comparison between Tool 1 and Tool 2

To support the measurement of the tool wear, the image of tool wear is captured using the MS3 Digital Microscope. After every cycle of the experiment, the tool is taken

out of the tool's holder and put under the microscope to be measured. Figure 15 shows the initial image of Tool 1. The initial length of the tool is 5.8035 mm before starting the machining process and Fig. 16 presents the tool wear image evolution from cycle 9 which is 3.8803 mm. Meanwhile, Fig. 17 shows the initial image of Tool 2 with the length is 6.0898 mm. After 9 cycles of experiment, the last tool length was reduced to 3.4945 mm with tool wear length as shown in Fig. 18.

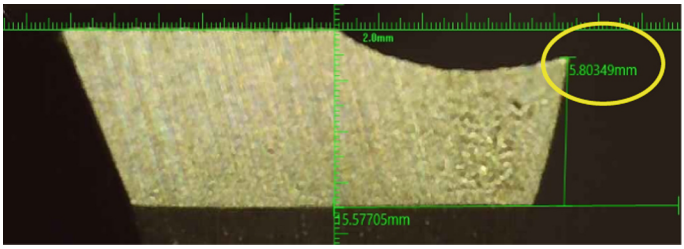


Fig. 15. Initial condition for Tool 1

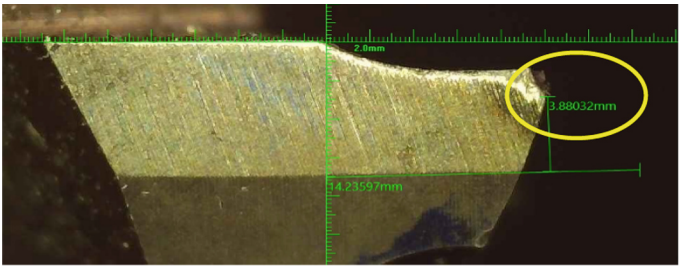


Fig. 16. Tool Wear Evolution for cycle 9 (Tool 1)

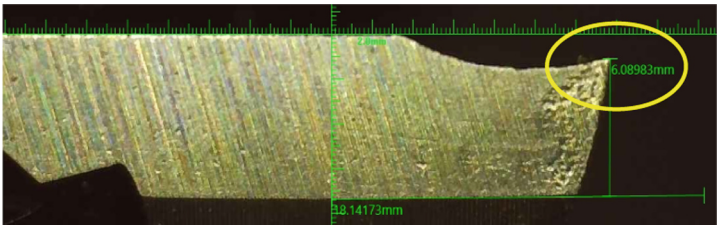


Fig. 17. Initial condition for Tool 2

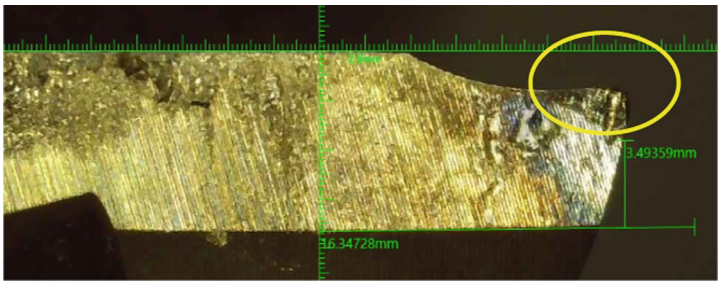
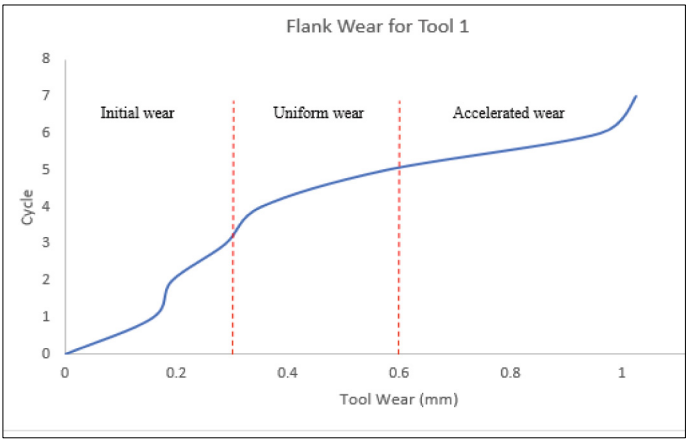


Fig. 18. Tool Wear Evolution for cycle 9 (Tool 2)

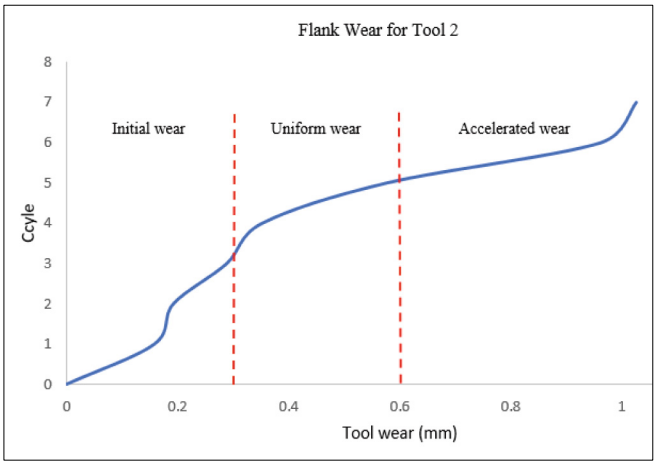
The progress of flank wear over the machining time is disclosed through three stages: the initial wear stage, uniform wear, and accelerated wear to evaluate the lifetime of a cutting tool. Table 3 discovered that it takes 6 cycles to obtain the 0.6 mm difference from the previous cycle for Tool 1. Based on Table 3, the Tool 1 wear value gradually increased from cycle 1 to cycle 9 starting from 5.80349 mm to 3.88032 mm. The tool wear becomes more visible as reaching the last cycle of the experiment. For Tool 2, a 0.6 mm difference also has been obtained on cycle 6 that exceeded to 0.9615 mm difference. The depth of cut value has become the key element to the tool wear breakage. The more depth of cut value, the faster it will get to reach 0.6 mm tool wear measurement. An illustration in Fig. 19 shows the difference in flank wear measurement on the tool wear image for Tool 1 and Tool 2.

Table 3. Tool Flank Wear for Tool 1 and Tool 2

	Tool 1	Tool 2
Cycle	Flank Wear (mm)	Flank Wear (mm)
Initial	0	0
1	0.0663	0.1582
2	0.1937	0.1924
3	0.2898	0.2885
4	0.3218	0.3523
5	0.5464	0.5769
6	0.6425	0.9615
7	0.7707	1.0253
8	1.0254	1.8269
9	1.6661	2.5963



a) Tool 1



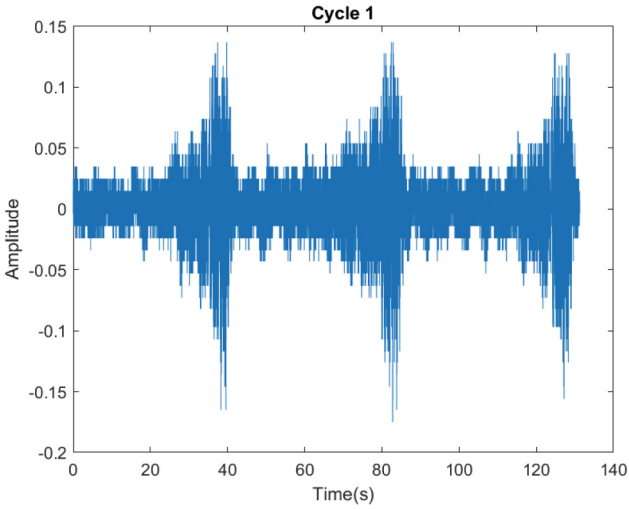
b) Tool 2

Fig. 19. (a) and (b): Cycle with flank wear measurement for Tool 1 and Tool 2

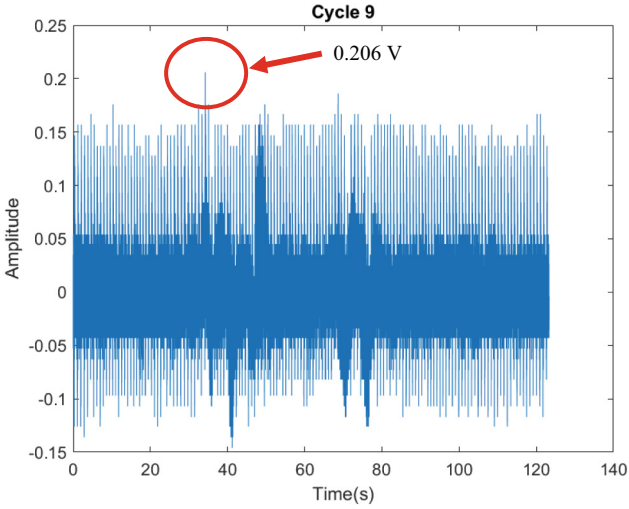
3.3 Time Domain Analysis

Time domain analysis of vibration signals is a method used to evaluate the characteristics of vibrations directly in the time domain, which involves examining how the vibration signal varies over time. This approach is fundamental for understanding the raw behavior of a system and is often the first step in vibration analysis before moving to more complex frequency domain methods. Time domain analysis refers to the real-time monitoring during the operation of the machine. Figure 20(a) shows the measured data obtained from the vibration sensor for Tool 1 and plots in the time domain graph for cycle 1 during the milling operation. From the pattern of the vibration signal, the trend of amplitude increases significantly with time taken during the experiment as shown in Fig. 20(b). This situation occurs due to the gradual wearing out of the tool during the milling process

leading to an increase in vibration amplitude due to higher cutting forces, irregular tool-workpiece interaction, and heat formed during the cutting process.



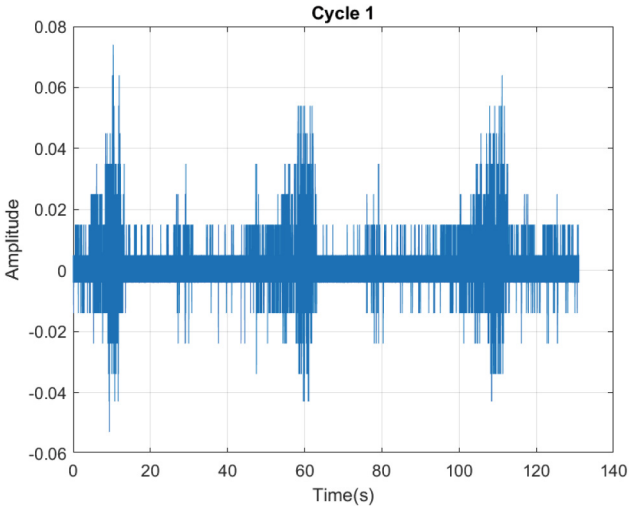
a) Tool 1 (for cycle 1)



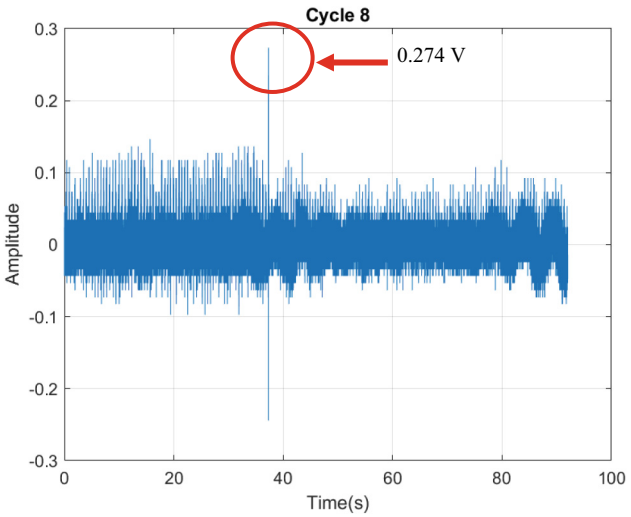
b) Tool 1(for cycle 9)

Fig. 20. (a) and (b): Time domain graph for vibration signal (Tool 1)

The comparison has been done for Tool 2 with different values of depth of cut as shown in Fig. 21(a) and (b) which represent cycle 1 and cycle 8 during the experiment. From the result, cycle 8 is the highest value of amplitude reached for Tool 2 about 0.274 V. Meanwhile the highest value of amplitude for Tool 1 is 0.206 V respectively.



a) Tool 2 (for cycle 1)

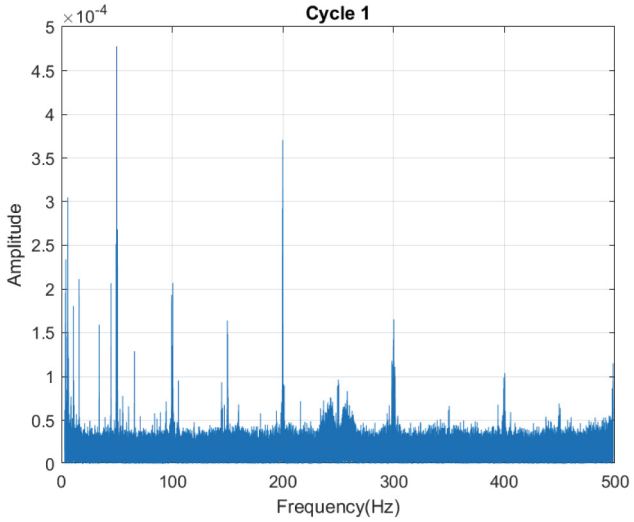


a) Tool 2 (for cycle 8)

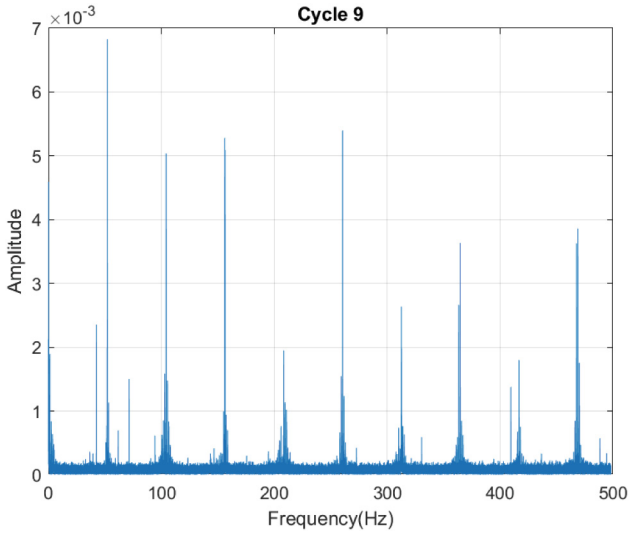
Fig. 21. (a) and (b) Time domain graph for vibration signal (Tool 2)

3.4 Frequency Domain Analysis (FFT)

Another method to analyze the tool vibration signal by applying a frequency domain method. In this research, the Fast Fourier transform (FFT) is used to convert time-domain data into frequency-domain data, making it practical for real-time signal processing and analysis. Figure 22(a) and Fig. 23(a) display the analysis of cutting tool vibration in the frequency domain for cycle 1, highlighting the frequency components present during the milling process. These figures illustrate the vibration signal's amplitude that is related to the frequencies. The resultant vibration amplitudes measured in the experiments, which showed an increase, are depicted in Fig. 22(b) and Fig. 23(b). This increase in amplitude is correlated with various factors observed during the milling process.



a) FFT for Tool 1 (cycle 1)

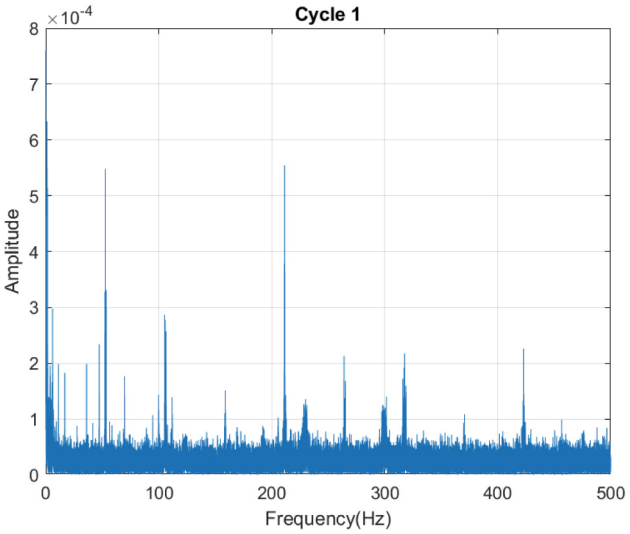


b) FFT for Tool 1 (cycle 9)

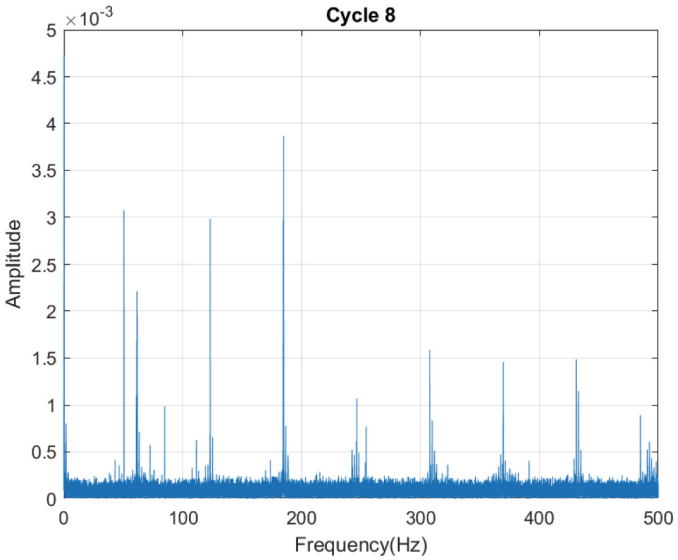
Fig. 22. (a) and (b): FFT graph for vibration signal (Tool 1)

As the depth of cut was increased, the amplitude of the vibration signals also rose, indicating a direct relationship between cutting depth and vibrational intensity. This is because a greater depth of cut increases the cutting forces, leading to more significant vibrations. Additionally, the duration of the cutting time impacted the amplitude of the vibration signals. Prolonged cutting times can lead to tool wear and thermal effects, both of which can contribute to increased vibrations. The graphs in Fig. 22(b) and Fig. 23(b) also reflect changes in amplitude due to noise generated during the milling process.

Noise in this context refers to unwanted or extraneous vibrations that can arise from various sources such as tool wear, material inconsistencies, and machine dynamics. These noise components can affect the overall vibration signal, leading to fluctuations in amplitude. By analyzing these graphs, it becomes evident how operational parameters, tool conditions, and external disturbances contribute to the vibration behavior of the cutting tool.



a) FFT for Tool 2 (cycle 1)



b) FFT for Tool 2 (cycle 8)

Fig. 23. (a) and (b): FFT graph for vibration signal (Tool 2)

4 Discussion

Fault detection on tool wear is essential for maintaining product quality, cost, and tool lifetime. This proactive approach also reduces unplanned downtime by optimizing tool usage and minimizing damage to workpieces. At present, diagnostic techniques rely on current, vibration, thermal analysis, and force measurements. This research focuses on vibration analysis to measure the tool wear fault detection.

This study focuses on vibration analysis for detecting tool wear, using a sensor-based method to analyze the vibration amplitude of two tools under varying parameters like speed, depth of cut, and feed rate. Tool wear measurements were captured using a digital microscope, showing a gradual reduction in tool length during the cutting process (Tool 1: 5.80349 mm to 3.88032 mm, Tool 2: 6.08983 mm to 3.49359 mm). Next is the data analysis study on time domain and frequency domain analysis, with Fast Fourier Transform (FFT) converting time-domain data into frequency-domain for real-time signal processing that compares with [13, 31] that focuses the frequency domain to analyze the cutting tool vibration. The findings demonstrate that increasing the depth of cut leads to higher cutting forces and vibration intensity, while extended cutting times increase tool wear and thermal effects. Noise from machine dynamics and tool wear affects vibration signals, indicating a need for noise filtering and further analysis of machining parameters to improve accuracy in tool wear prediction.

The proposed vibration-based tool wear detection method improves real-time monitoring that controls product quality, cost savings, and tool lifetime. Integrating both time domain and frequency domain analysis with Fast Fourier Transform (FFT) offers a more detailed understanding of tool wear under different conditions, such as speed and depth of cut. This approach reduces unexpected downtime and maximizes tool efficiency, while also accounting for external factors like noise from machine dynamics. The method shows real-time monitoring systems have a potential for improvement through noise filtering and further analysis of machining parameters.

5 Conclusion

In conclusion, this research presents an enhancement method for fault diagnosis and type classification for tool wear. The vibration voltage amplitude is obtained from the sensor and sends the data through the DAQ device. It has focused on three main areas: the measurement of tool wear image, the measurement data from the vibration sensor, and the processing of that sensor data to determine a tool wear condition. The experiment result shows that the amplitude of the vibration signal is affected by the variations in depth of cut (0.5 mm and 1.0 mm). The vibration signal for Tool 1 is achieved 0.206 V and Tool 2 is achieved 0.274 V respectively. During the experiment, the tool image is captured using high-resolution cameras to obtain a view of the flank surface. Advanced image processing software is then employed to measure the length of flank wear, which is defined as the distance along the flank surface where significant wear has occurred. Initial wear is documented before the experiment starts, and periodic images are captured to monitor the progression of wear over time. The result of the tool wear length image measurement for Tool 1 was 5.80349 mm reduced to 3.88032 mm and for Tool 2 was 6.08983 mm

reduced to 3.49359 mm during the cycle 9 experiment. Each tool reaches an accelerated region during cycle 6 which is 0.6425 mm and 0.9615 mm flank wear measurement. For future improvement, the data analysis for tool wear measurement can be significantly enhanced by incorporating artificial intelligence (AI) methods for the classification of tool wear. AI techniques, such as machine learning and deep learning algorithms, can analyze large datasets more efficiently and accurately than traditional methods. By training models on historical wear data, these algorithms can learn to recognize patterns and make predictions about tool wear under various operating conditions.

Acknowledgment. This research is fully supported by UC TATI short-term research grant (STG-phase 1/GPJP2023). The authors fully acknowledged the support of the University Malaysia Pahang and University College TATI for their support of this work.

References

1. Ghule G, Ambhore NH, Chinchanikar S (2018) Tool condition monitoring using vibration signals during hard turning: a review. SSRN Electron J. https://doi.org/10.2139/ssrn.3101977
2. Zhu K, Yu X (2017) The monitoring of micro milling tool wear conditions by wear area estimation. Mech Syst Signal Process 93:80–91. https://doi.org/10.1016/j.ymssp.2017.02.004
3. Peng R, Pang H, Jiang H, Hu Y (2020) Study of tool wear monitoring using machine vision 54:259–270. https://doi.org/10.3103/S0146411620030062
4. Gomes MC, Brito LC, Bacci da Silva M, Viana Duarte MA (2021) Tool wear monitoring in micromilling using Support Vector Machine with vibration and sound sensors. Precis Eng 67:137–151. https://doi.org/10.1016/j.precisioneng.2020.09.025
5. Li W, Liu T (2019) Time varying and condition adaptive hidden Markov model for tool wear state estimation and remaining useful life prediction in micro-milling. Mech Syst Signal Process 131:689–702. https://doi.org/10.1016/j.ymssp.2019.06.021
6. Upase R, Ambhore N (2020) Experimental investigation of tool wear using vibration signals: an ANN approach. Mater Today Proc 24:1365–1375. https://doi.org/10.1016/j.matpr.2020.04.454
7. Twardowski P, Tabaszewski M, Felusiak-Czyryca A, Kieruj P, Wiciak–pikuła M, Czyżycki J (2022) Identification of tool wear during cast iron drilling using machine learning methods. Adv Sci Technol Res J 16:126–139. https://doi.org/10.12913/22998624/155985
8. Zhou Y, Liu C, Yu X, Liu B, Quan Y Tool wear mechanism, monitoring and remaining useful life (RUL) technology based on big data: a review. SN Appl Sci 4 (2022). https://doi.org/10.1007/s42452-022-05114-9
9. Xu G, Liu M, Jiang Z, Shen W, Huang C (2020) Online fault diagnosis method based on transfer convolutional neural networks. IEEE Trans Instrum Meas 69:509–520. https://doi.org/10.1109/TIM.2019.2902003
10. Madhusudana CK, Gangadhar N, Kumar H, Narendranath S (2018) Use of discrete wavelet features and support vector machine for fault diagnosis of face milling tool 12:111–127. https://doi.org/10.3970/sdhm.2018.01262
11. Li X, Deng, Y., Li, X (2021) Application of multisensor information fusion technology in the measurement of dynamic machining errors of computer numerical control (CNC) machine tools. J Sensors 2021. https://doi.org/10.1155/2021/6918496
12. Lu N, Xu B, Chen Z, Ren D, Yang Y, Bian P (2020) Research simulation on fault diagnosis of CNC machine tool based on fuzzy petri net. Springer Nat. Singapore Pte Ltd. 2020. CISC 2019 41–49. https://doi.org/10.1007/978-981-32-9686-2_6

13. Md Tahir NH, Rizal M, Mansor WNW, Ghani JA, Nuawi MZ, Haron CHC (2017) The correlation between tool wear and vibration signals from piezoelectric in end milling of AISI P20+NI. J Mech Eng SI 3:49–62
14. Díaz-Saldaña G, Osornio-Ríos RA, Zamudio-Ramírez I, Cruz-Albarrán IA, Trejo-Hernández M, Antonino-Daviu JA (2023) Methodology for tool wear detection in CNC machines based on fusion flux current of motor and image workpieces. Machines 11. https://doi.org/10.3390/machines11040480
15. Wszołek G, Czop P, Słoniewski J, Dogrusoz H (2020) Vibration monitoring of CNC machinery using mems sensors. J Vibroengineering 22:735–750. https://doi.org/10.21595/jve.2019.20788
16. Wang Q, Wang H, Hou L, Yi S (2021) Overview of tool wear monitoring methods based on convolutional neural network. Appl Sci 11 . https://doi.org/10.3390/app112412041
17. Sridhar AV, Prasad BS, Mouli KVVNRC (2021) Evaluation of tool performance and wear through vibration signature analysis in the drilling of IS3048 steel. J Eng Appl Sci 68:1–12. https://doi.org/10.1186/s44147-021-00036-6
18. Adipovich MK, Anvarovich MR (2020) Diagnostic of the state of the cutting tool of metal-cutting machines using bidirectional recurrent neural networks with a long short-short-term memory 1093–1097. https://doi.org/10.1109/ICARCV50220.2020.9305486
19. Martínez-Arellano G, Terrazas G, Ratchev S (2019) Tool wear classification using time series imaging and deep learning. Int J Adv Manuf Technol 104:3647–3662. https://doi.org/10.1007/s00170-019-04090-6
20. Chen Y, Bian R, Ding W (2019) A fault diagnosis method of CNC machine tool spindle based on deep transfer learning. Comput Inf Anal Intell Syst (CIAIS 2019) 136–140. https://doi.org/10.25236/ciais.2019.027
21. Wiciak-Pikuła M, Felusiak-Czyryca A, Twardowski P (2020) Tool wear prediction based on artificial neural network during aluminum matrix composite milling. Sensors (Switzerland) 20:1–18. https://doi.org/10.3390/s20205798
22. Zhang XY, Lu X, Wang S, Wang W, Li WD (2018) A multi-sensor based online tool condition monitoring system for milling process. Procedia CIRP 72:1136–1141. https://doi.org/10.1016/j.procir.2018.03.092
23. Jáuregui JC, Member S, Reséndiz JR, Member S (2018) Frequency and time-frequency analysis of cutting force and vibration signals for tool condition monitoring 6. https://doi.org/10.1109/ACCESS.2018.2797003
24. Sushant J, Manas S, Abhishek P, Amey O, Ajaykumar U: Analysis of tool wear in milling using machine learning technique. Int Res J Eng Technol 1602–1609 (2023)
25. Nouioua M, Bouhalais ML (2021) Vibration-based tool wear monitoring using artificial neural networks fed by spectral centroid indicator and RMS of CEEMDAN modes. Int J Adv Manuf Technol 115:3149–3161. https://doi.org/10.1007/s00170-021-07376-w
26. Lee WK, Abdullah MD, Ong P, Abdullah H, Teo WK (2021) Prediction of flank wear and surface roughness by recurrent neural network in turning process. J Adv Manuf Technol 15:55–67
27. Kamdani K, Hasan S, Ahmad Ashaary AFI, Lajis MA, Rahim EA (2019) Study on tool wear and wear mechanisms of end milling nickel-based alloy. J Tribol 21:82–92
28. Chen SH, Luo ZR (2020) Study of using cutting chip color to the tool wear prediction. Int J Adv Manuf Technol 109:823–839. https://doi.org/10.1007/s00170-020-05354-2
29. Stavropoulos P, Papacharalampopoulos A, Vasiliadis E, Chryssolouris G (2016) Tool wear predictability estimation in milling based on multi-sensorial data. Int J Adv Manuf Technol 82:509–521. https://doi.org/10.1007/s00170-015-7317-6

30. Huang PM, Lee CH (2021) Estimation of tool wear and surface roughness development using deep learning and sensors fusion. Sensors 21. https://doi.org/10.3390/s21165338
31. Rad JS, Zhang Y, Aghazadeh F, Chen ZC (2014) A study on tool wear monitoring using time-frequency transformation techniques. In: Proceedings of the 2014 international conference on innovative design and manufacturing (ICIDM) 2014, vol 17, pp. 342–347. https://doi.org/10.1109/IDAM.2014.6912718

PID Control System Design to Improve the Performance of Negative Pressure Wound Therapy

I. Dewa Gede Hari Wisana(✉), Priyambada Cahya Nugraha, Levana Forra Wakidi, Danu Rogo, and Riqqah Dewiningrum

Department of Electromedical Engineering, Health Polytechnic Ministry of Health Surabaya, Surabaya, Indonesia

hariwisana@yahoo.com

Abstract. Treatment of infected and colonized diabetic ulcers requires a multidisciplinary approach, including wound cleansing to remove exudate, reduce bacterial load, and promote granulation tissue formation. Negative pressure therapy (NPWT) can help accelerate the healing of diabetic ulcers by suctioning exudate through a tube connected to a suction device, thereby accelerating the process of tissue regeneration. This study aims to increase local content in affordable and effective diabetic ulcer therapy methods by developing a simple and low-cost NPWT prototype. This prototype uses an ESP32 microcontroller as the programming processor, which includes PID (Proportional Integral Derivative) control, XGZP101DB1R sensor reading, therapy timer, as well as three modes of operation namely continuous, intermittent, and dynamic, alarm system and displayed on a Nextion TFT LCD. This tool uses a rechargeable battery power source. There is also a battery level indicator and charging status. Its lightweight and moderate design facilitates mobility. This research develops a negative pressure control system for vacuum DC motors using PID to improve the precision of instrumentation systems with feedback. By determining the Kp, Ki, and Kd values, pressure regulation is expected to be performed more accurately, providing additional information to the healthcare team. The prototype can read negative pressure from 0 to −150 mmHg and has been calibrated before testing to minimize errors. Tests were carried out using a Digital Pressure Meter (DPM). Negative pressure settings of −50, −75, −100, −125, and −150 mmHg were tested and performed in all three modes: continuous, intermittent, and dynamic. The test results show that the Negative Pressure Wound Therapy (NPWT) prototype based on PID control can achieve negative pressure according to the set point with the largest average error of 0%, 0%, and 1.13% and the lowest error of 0%, 0%, and 0.31% in continuous, intermittent, and dynamic modes, respectively. The highest error rate in pressure measurement using a Digital Pressure Meter (DPM) is 2.33%, and the lowest is 0.15% in all modes. In conclusion, the ESP32 system and PID control on the NPWT therapy device are effective in helping to accelerate the wound healing process and diabetic ulcer recovery with minimal error rates.

Keywords: Diabetic ulcer · XGZP101DB1R sensor · ESP32 · continuous · intermittent and dynamic

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 361–377, 2026.
https://doi.org/10.1007/978-981-95-9233-3_26

1 Introduction

Diabetes is a multifaceted metabolic disease that affects over 340 million people globally, with approximately 20% of them suffering from diabetic wounds [1]. Approximately 50–70% of all limb amputations are due to diabetic wounds, and it is reported that every 30 s, one foot is amputated due to diabetic wounds worldwide [2]. Wounds can be defined as disturbances in tissue continuity with a variety of causes and can have a significant impact on individuals if not managed properly. The healthcare system allocates human and financial resources to create an optimal environment to support wound healing [3]. Therefore, wound healing has always been a focus of research across various medical fields, presenting significant challenges and economic burdens [4, 5]. Actions are needed to prevent bacterial infections and accelerate the wound-healing process [6–9]. Wound healing is a dynamic process that restores anatomical structure and similar functions. Effective wound management requires fast and complete healing without infection or sepsis. Acute wounds generally heal without issues, but age remains a primary concern [10, 11]. Skin damage caused by trauma, burns, harmful chemicals, genetic disorders, or systemic diseases poses a serious threat to life and health. The wound-healing process is complex and involves stages such as hemostasis, inflammation, granulation, and remodeling [12, 13]. Treatment of infected and colonized wounds requires a multidisciplinary approach that includes different types of treatments, including surgical debridement to remove necrotic and damaged tissue. In addition, systemic treatment with antibiotics is necessary, along with wound cleaning to remove exudate, reduce bacterial load, and promote the formation of granulating tissue at the wound site. Furthermore, the closure or protection of the wound is carried out safely according to the complexity of the wound [14]. Although many physiological healing processes have been described in the literature for centuries, recent research and trials have identified the optimal environment to enhance the healing process. Several factors, such as necrotic tissue, foreign bodies, and the risk of infection, must be addressed to ensure better wound closure [3].

Negative pressure wound therapy (NPWT) is used to improve the wound healing process through several mechanisms, such as changes in blood flow around the wound, mechanical deformation of the wound peripheral tissue, and excretion of excess fluid and debris [15]. NPWT is a method used to treat superficial wounds and drain deep wounds. This therapy helps remove necrotic tissue and exudate, stimulates the formation of granulation tissue, reduces bacterial infection, and accelerates wound healing [16]. NPWT is a mechanical device that uses a tube connected to a suction unit to create negative pressure between the wound and the external environment, which helps remove wound exudate and accelerate healing [17, 18]. This therapy has been successfully used for various types of wounds, such as traumatic wounds, pressure ulcers, leg ulcers, surgical wound dehiscence, necrotic wounds, and venous ulcers. In some cases, NPWT can simplify the surgical procedures required for treatment [19, 20].

NPWT has been clinically proven to accelerate the healing process of open wounds by three times or more, as well as aid in total recovery with less new tissue formation [21]. The NPWT system adopts 'continuous' and 'intermittent' modes. In continuous mode, a negative pressure of −125 mmHg is run continuously, while in intermittent mode, a negative pressure of −25 mmHg is turned on for 5 min, followed by a 2-min rest phase at 0 mmHg. Intermittent mode has been shown to increase perfusion and formation of granulated tissue around the wound compared to the continuous mode. Although effective in wound healing, the intermittent mode is often avoided in clinical use because it causes pain that occurs every few minutes during the initiation phase of the system to reach a pressure of 125 mmHg [22].

In 2011, Haijun NIU et al., made a study "Development of a Device for Negative Pressure Wound Therapy based on ARM9 Embedded System", using an embedded system with a kernel arm, that can operate with two different negative pressure sources (pump and clinical pressure), and the vacuum level can be set using the PWM control method. However, in this study using the PWM method, there has been no research conducted using the PID (Proportional-Integral-Derivative) control method [23]. Furthermore, in 2017, Hou Tan Jia et al., made research "Incorporating Fuzzy Logic Into An Adaptive Negative Pressure Wound Therapy Device". This tool uses stable negative pressure control with a fuzzy logic controller (FLC). This NPWT system can provide negative pressure from 0 to 200 mmHg, which can be controlled and still uses continuous mode. However, in this research, the NPWT tool designed only uses one mode, namely the continuous mode [24]. The next trace in 2020, made by Fikri Fahriansyah et al., with the title "A Low Cost Negative Pressure Wound Therapy Based on Arduino" on the tool utilizes MPXV4115VC6U and MPXV5050VC6T1 pressure sensors in the pressure range from 0 to −350 mmHg. The Arduino microcontroller is used to process data from the sensor, and the results are displayed on the liquid crystal display (LCD), for the mode it still uses the continuous mode in the therapy process and is equipped with a time setting. However, only one mode is used, namely continuous mode in the therapy process [25]. Then in 2022, continued by Auria Ulul Amanah with the title "Negative Pressure Wound Therapy Equipped with 2 Modes". This tool uses 2 modes, namely continuous mode and intermittent mode, the negative pressure is generated from the vacuum motor which uses a MPXV4115VC6U sensor to read the negative pressure at pressures between −50 to −150 mmHg with increments every −25 mmHg in continuous mode while using pressures of 0 and −125 mmHg in intermittent mode. Using an Arduino microcontroller processor for data processing is then displayed on the LCD in the form of TFT Nextion. However, this study only uses two modes and no additional dynamic modes [26]. In 2023, Angga Davida et al. developed software to control the NPWT using an Arduino UNO microcontroller to execute the output of the MPXV4115VC6U pressure sensor, timer, and PID controller. In addition, it uses a 16 x 2 LCD, L298N motor driver, 5V buzzer, and AC to DC power adapter. With the therapy modes provided, namely continuous and intermittent with minimum and maximum pressures of −75 to −125 mmHg. However, this research only provides two therapy modes, the maximum pressure is −125 mmHg, has not used a TFT LCD to make the display more attractive, and has not used a battery so it does not allow for portability [27].

The development of Negative Pressure Wound Therapy (NPWT) technology continues to improve effectiveness, patient comfort, and minimize potential risks. The development of sensors helps in more accurate settings and provides more information to healthcare teams. Pressure control includes the ability to regulate vacuum pressure with more precision with PWM settings or a Fuzzy Logic system [23, 24]. Improvement of healing ability and acceleration of the process in the form of setting modes used, namely continuous, intermittent, and dynamic modes [28, 29]. Although many studies have been conducted to design NPWTs, there are still few NPWTs designed by providing three modes, namely continuous, intermittent, and dynamic with a PID-based controller. In addition, the NPWT designed in previous studies has not used batteries so the actions performed must be near a power source and cannot increase mobility and comfort for users.

Based on the above identification, the author wants to develop a negative pressure control system on DC vacuum motors using a PID (Proportional Integral Derivative controller) system to determine the precision of an instrumentation system with the characteristics of feedback on the system in finding the values of Kp, Ki, and Kd [30]. NPWT therapy uses continuous, intermittent, and dynamic modes. Continuous mode in the constant negative pressure configuration, intermittent mode in the constant pressure configuration turns on and off repeatedly, while dynamic mode is the same as intermittent mode but the on and off pressure values can be set according to the setting [31]. Then it will be displayed on the LCD. With this description, the author will make a study entitled "PID System Control Design to Improve the Performance of Negative Pressure Wound Therapy" as a development of previous research, hopefully the development of the tool can provide a reference in research on negative pressure therapy tools.

2 Materials and Method

This research was conducted at the Polytechnic of the Ministry of Health Surabaya. Under the Department of Electromedical Technology. Negative pressure wound therapy as a therapy tool for this study. The research design used in the development of the module is a pre-experimental design with a type of research after only design. In this design, the author sees the results of the practice, namely with the modes in the type of therapy, namely continuous, intermittent, and dynamic modes. The focus of this design is to conduct an in-depth analysis of the PID control variables in various settings.

This study uses the XGZP101DB1R sensor as input to display the results of real negative pressure readings and negative pressure settings. The data obtained from the sensor readings XGZP101DB1R was processed using the ESP32 microcontroller. In this study, the independent variable is the negative pressure on the chamber canister. The bound variable is the XGZP101DB1R sensor used to detect the negative pressure reading and the vacuum pump is used as a series of suction pressure sources. The control variable in the creation of this module is ESP32.

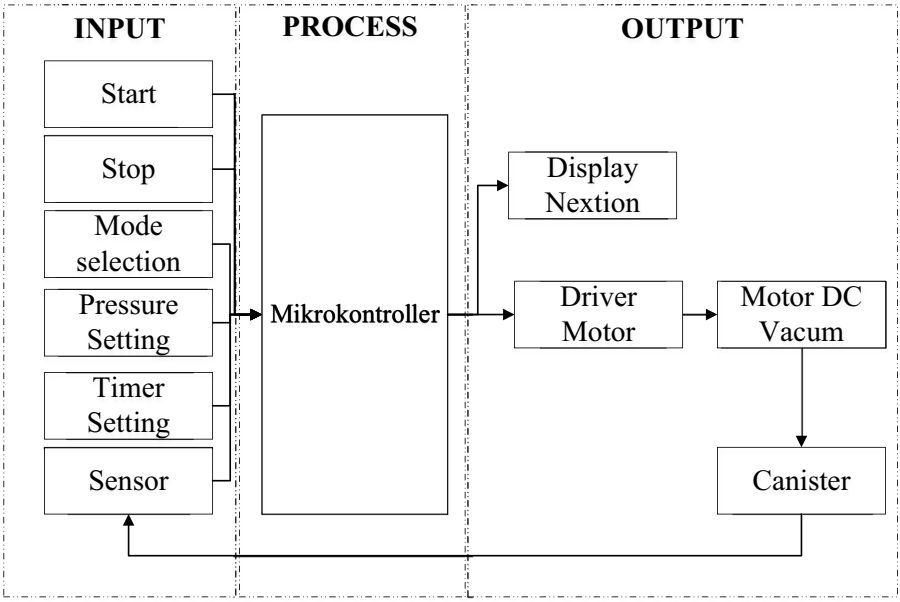


Fig. 1. System block diagram on pressure wound therapy negative device research with PID control

This paragraph will explain Fig. 1 illustrates the system block diagram, which consists of 3 main parts: input, process, and output. In the input section, there is an input from the XGZP101DB1R sensor, where the XGZP101DB1R sensor will detect the real negative pressure when the device is turned on. Then, in the input section, there is a negative pressure setting, where the desired negative pressure setting value is entered through the Nextion screen with a pressure range of 0 to 150 mmHg for the desired pressure. After the negative pressure setting is entered, set the timer for the therapy process, press the enter and start button, and then the ESP32 microcontroller will process the negative pressure setting data using the PID control. Then the pump driver will control the vacuum pump circuit, next, the system will work so that the negative pressure sensor functions to identify the negative pressure that is active in the reservoir until it reaches the setpoint pressure according to the desired pressure. When the pressure has reached the setpoint, the vacuum pump will lower the pressure and keep the pressure stable. Then the Nextion LCD will display the setting pressure accompanied by the reading pressure.

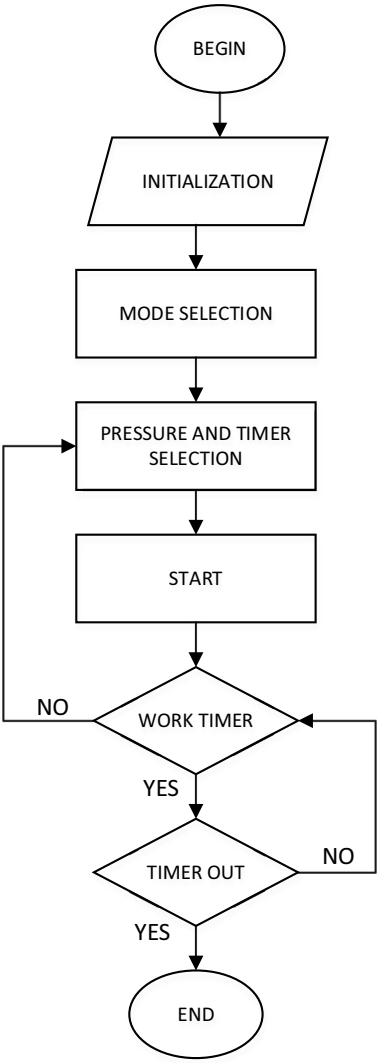


Fig. 2. Flow diagram of pressure wound therapy negative tool module with PID control

The module flow diagram in Fig. 2 explains that when the module is installed in the right position, press the ON button after the module is turned on, and the tool will initialize the data. Then select the mode you want. When the continuous mode is in continuous mode, the negative pressure setting and the timer. When in intermittent mode, the negative pressure setting and its timer, when in dynamic mode, the upper threshold pressure setting, the lower threshold pressure setting, and its timer. For intermittent and dynamic modes, one cycle timer is on and the timer is off [15]. The start button is to start the therapy and the stop button is to end the therapy. And when the therapy is finished,

it will automatically return to the initial setting, the system will be in a standby position and ready to do therapy again.

2.1 Data Analysis

The research was carried out by measuring the negative pressure wound therapy device module using PID control compared to a digital pressure meter. Testing the negative pressure wound therapy module using PID control compared to the digital pressure meter was carried out repeatedly 5 times with 5 measurement points namely −50, −75, −100, −125, and −150 mmHg. Testing was also carried out for 5 min using three modes, namely continuous, intermittent, and dynamic modes at 5 measurement points −50, −75, −100, −125, and −150 mmHg to see the PID control response to the system. A PID controller is composed of three elements: Proportional (P), Integral (I), and Derivative (D). Different combinations of these components create various types of PID controllers, each serving specific purposes. PID controllers are used in many fields, and their requirements can vary depending on the system being controlled. For instance, PI controllers emphasize quick response and aim to reduce or eliminate steady-state errors, while PD controllers provide faster response and stability. Generally, when designing a PID controller, it must satisfy certain conditions: it must meet performance targets, it should be based on a known or accessible process, and it must account for the available computational resources. The PID control system is a second-order linear controller [32]. The basic equation of the PID algorithm is shown in Eq. (1) as follows:

$$(t) = K_p\left[e(t) + \frac{1}{T_i}\int_0^t e(t)dt + T_d\frac{de(t)}{dt}\right] \tag{1}$$

where u(t) is the output of the control system, e(t) is the input of the control system, which is the error between the target value and the current value, Kp is the scale factor of the control system. Ti denotes the integration time of the control system, then Td is the differentiation time of the control system. To obtain the accuracy value on the negative device module, pressure wound therapy uses PID control compared to a digital pressure meter. Measurements have been carried out 5 times. From 5 repeated measurements, it is necessary to calculate the average. By applying Eq. (2) The average is used to determine the average value of the measurement:

$$\left(\overline{X}\right) = \frac{\Sigma Xi}{n} \tag{2}$$

by using arithmetic averages, it means that it is a number that is produced by dividing the total value of the data by the amount of data in the set. Average is the value or result of dividing the amount of data taken or measured by the total data taken or the number of measurements. Where x represents the mean value for xi is all the data that exists where n is the amount of data that exists. A standard deviation is a number that shows how much variance there is in a set of data or standard deviation from the mean. Equation (3) can be used to display standard deviation (SD) formulas:

$$X = \sqrt{\frac{(Xi - X)^2}{n-1}} \tag{3}$$

where xi is the percentage of the value in question, x indicates the average of the measurement results, and n indicates the total number of measurements. The doubts that may arise on each measurement result are called uncertainties (UAs). In Eq. (4), the uncertainty formula is displayed:

$$UA = \frac{SD}{\sqrt{n}} \tag{4}$$

where UA indicates the overall degree of uncertainty of the measurement, SD depicts the standard deviation of the measurement as a result and indicates the total number of measurements. System errors are displayed by % errors. Smaller numbers The difference between the averages of each data set is an error. Errors may indicate how the model or design deviates from the norm. Equation (5) display error formulas.

$$ERROR = \frac{(Xn - X)}{Xn} x100\% \tag{5}$$

where Xn represents the value measured by the calibrator. The specified value of the design is x. An error is the difference between the average value of each data. This study shows that the accuracy value is still within the tolerance limit.

3 Result

After the design of the negative pressure wound therapy device module using PID control is completed, the next step is to collect data. This research is to collect data. Data collection was carried out by measuring 5 times on the module with a digital pressure meter. During the data retrieval process, measurement recording is done on Microsoft Excel.

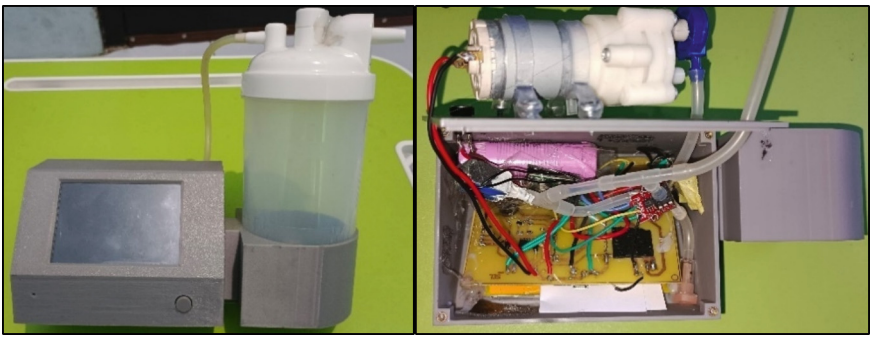


Fig. 3. Range of components used in tools

Figure 3 shows the electronic series of the negative pressure wound therapy device module using PID control by utilizing several components. From the results of the design of the negative pressure wound therapy module that uses PID control, this study uses the AC voltage of 220 V from PLN as the main source of electricity used to supply the voltage at the 12 V/10 A power supply. In addition, this module consists of several important components, namely the adapter, battery, ESP32, and sensor module XGZP101DB1R the control circuit in this module is carefully implemented. In addition, in the vacuum pump control circuit, there are components such as vacuum pumps, solenoids, pump drivers, buzzers, and XGZP101DB1R sensors. All of these components are interconnected to create the desired function.

In addition to these main components, this module is also equipped with various additional features. One of them is the 3.2-inch Nextion LCD, which is used to display the real negative pressure and setting pressure. There is also an on switch that allows the user to turn on the appliance.

Table 1. Measurement results on calibrators and modules in continuous mode

Setting	Measurement tools	Mean	SD	Error (%)
−50	DPM	−43.3	0.35	0.82
	Module	−50	0.00	0.00
−75	DPM	−72.2	0.35	0.49
	Module	−75	0.00	0.00
−100	DPM	−101.2	0.27	0.27
	Module	−100	0.00	0.00
−125	DPM	−125.8	0.27	0.22
	Module	−125	0.00	0.00
−150	DPM	−150.3	0.27	0.18
	Module	−150	0.00	0.00

Table 1 results shows the comparison between the measurement results of the digital pressure meter and the module. For readings on digital pressure meters, the lowest error value is 0.18% of the −150 mmHg setting and the highest error value is 0.82% of the −50 mmHg setting. For readings on the module, the lowest error value is 0% and the highest error value is 0%.

Figure 4 is the result of negative pressure plotting using a PID system with pressure settings of −50, −75, −100, −125, and −150 mmHg. At a pressure setting of −50 mmHg, negative pressure is reached and stabilized at 15 s. At a pressure setting of −75 mmHg, negative pressure is reached and stabilized at 13.5 s. At a pressure setting of −100 mmHg, negative pressure is reached and stabilized at 13.8 s. At a pressure setting of −125 mmHg, negative pressure is reached and stabilized at 6 s. At a pressure setting of −150 mmHg, negative pressure is reached and stabilized at 7.8 s. From the

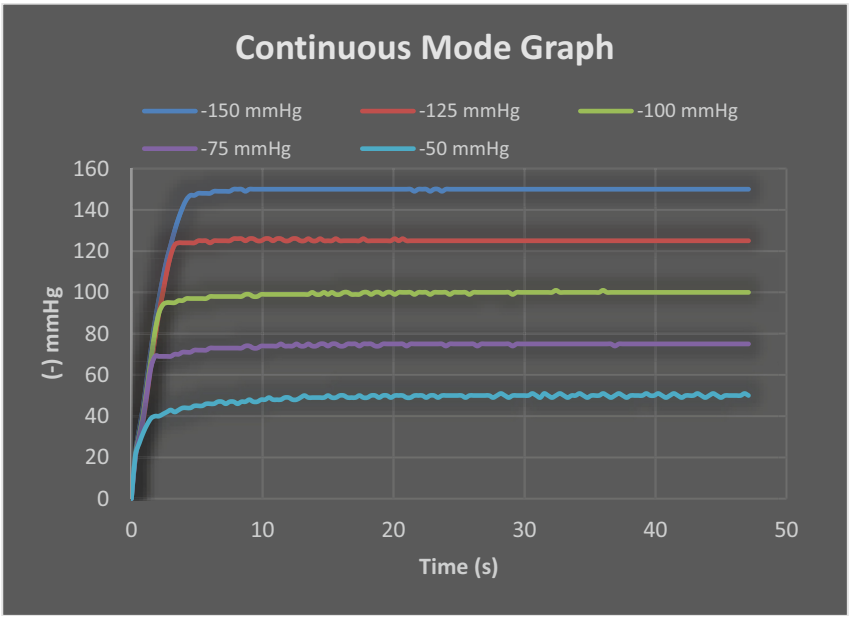


Fig. 4. Mode chart on the pressure wound therapy negative device using PID control in continuous mode

measurement error graph on the module, it can be explained that the smaller the error value, the better and more accurate the calibration results will be. On the other hand, the greater the error value in the measurement results, the less accurate the measurement will be.

Table 2. Measurement results on calibrators and modules in intermittent mode

Setting	Measurement Tools	Mean	SD	Error (%)
−50	DPM	−43	0.35	0.82
	Module	−50	0.00	0.00
−75	DPM	−72.5	0.35	0.49
	Module	−75	0.00	0.00
−100	DPM	−101.7	0.27	0.27
	Module	−100	0.00	0.00
−125	DPM	−126.3	0.27	0.22
	Module	−125	0.00	0.00
−150	DPM	−150.4	0.22	0.15
	Module	−150	0.00	0.00

Table 2 results shows the comparison between the measurement results of the digital pressure meter and the module. For readings on digital pressure meters, the lowest error value is 0.15% of the −150 mmHg setting and the highest error value is 0.82% of the −50 mmHg setting. For readings on the module, the lowest error value is 0% and the highest error value is 0%.

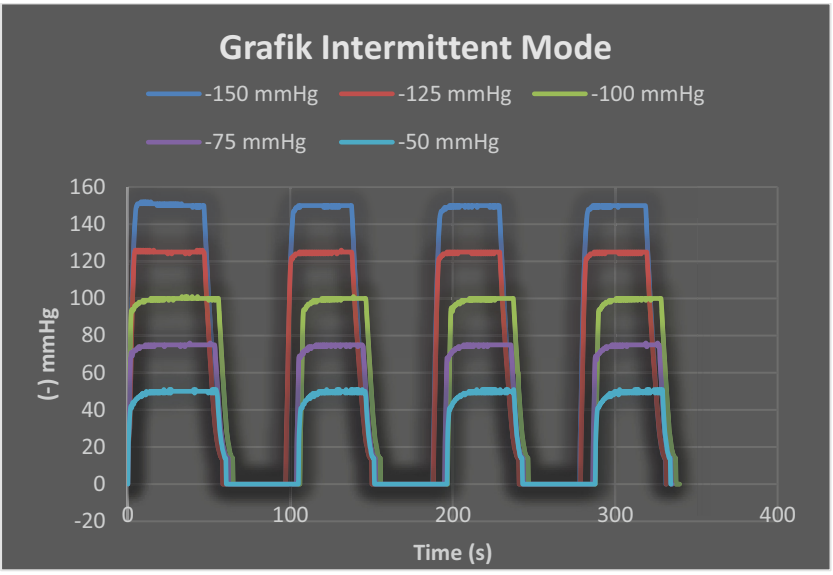


Fig. 5. Mode chart on the pressure wound therapy negative device using PID control in intermittent mode

Figure 5 is the result of negative pressure plotting using the PID system with pressure settings of −50, −75, −100, −125, and −150 mmHg, setting a rise time of 1 min and a fall time of 1 min. At a pressure setting of −50 mmHg, the first cycle of negative pressure is reached and stabilized at 14.4 s. At a pressure setting of −75 mmHg, negative pressure was reached and stabilized at 12.3 s. At a pressure setting of −100 mmHg, negative pressure was reached and stabilized at a time of 14.1 s. At a pressure setting of −125 mmHg, negative pressure was reached and stabilized at a time of 4.8 s. At a setting pressure of −150 mmHg, negative pressure was reached and stabilized at 5.7 s.

Table 3 results show the comparison between the measurement results of the digital pressure meter and the module. For readings on digital pressure meters, the lowest error value is 0.27% of the setting, −100 mmHg, and the highest error value is 2.33% of the −50 mmHg setting. For readings on the module, the lowest error value is 0.31% of the setting, −150 mmHg and the highest error value is 1.13% of the setting, −75 mmHg.

Figure 6 is the result of negative pressure plotting using the PID system with maximum pressure settings of −75, −100, −125, and −150 mmHg, minimum pressure of −50 mmHg, rise time of 1 min, and fall time of 1 min. At a pressure setting of −150 mmHg, the first cycle of negative pressure was achieved with a maximum pressure

Table 3. Measurement results on calibrators and modules in dynamic mode

Setting	Measurement Tools	Mean	SD	Error (%)
−50	DPM	−43.9	1.02	2.33
	Module	−50.4	0.55	1.09
−75	DPM	−71.9	0.65	0.91
	Module	−74.2	0.84	1.13
−100	DPM	−99.7	0.27	0.27
	Module	−97.8	0.45	0.46
−125	DPM	−121	0.61	0.51
	Module	−120.2	0.45	0.37
−150	DPM	−144.4	0.42	0.29
	Module	−144.8	0.45	0.31

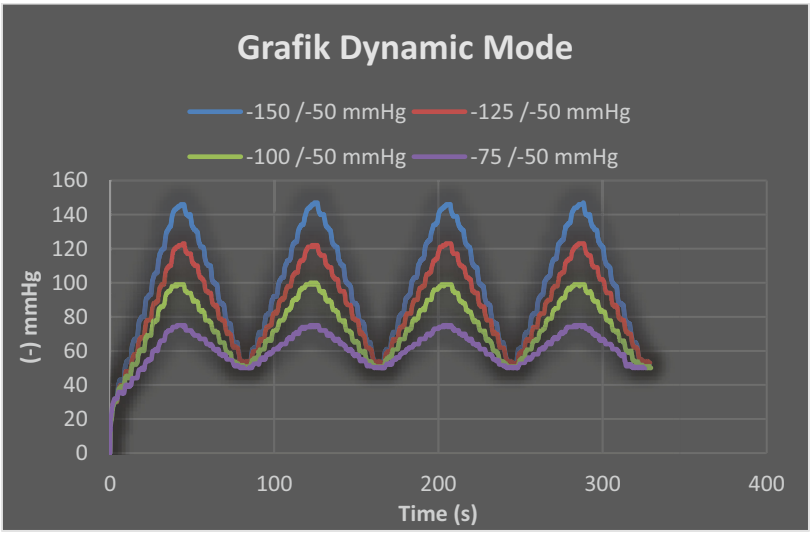


Fig. 6. Mode chart on the pressure wound therapy negative device using pid control in dynamic mode

of −146 mmHg. At a pressure setting of −125 mmHg, the first cycle of negative pressure was reached with a maximum pressure of −124 mmHg. At a pressure setting of −100 mmHg, the first cycle of negative pressure is reached with a maximum pressure of −99 mmHg. At a pressure setting of −75 mmHg, the first cycle of negative pressure was reached with a maximum pressure of −75 mmHg.

Table 4. T-test on the continuous mode of Negative Pressure Wound Therapy using PID Control

	Variable 1	*Variable 2*
Mean	100	98.52
Variance	1562.5	1788.467
Observations	5	5
Pooled Variance	1675.4835	
Hypothesized Mean Difference	0	
df	8	
t Stat	0.057169138	
P(T <= t) one-tail	0.477906289	
t Critical one-tail	1.859548038	
P(T <= t) two-tail	0.955812578	
t Critical two-tail	2.306004135	

Table 4, is a t-test of negative pressure wound therapy using PID control. The t-test is used to compare the means of two groups of data and determine whether the difference between the means is statistically significant or merely coincidental. From the t-test results, it can be seen that the null hypothesis (H0) states that there is no difference in mean DPM between the modules, while the alternative hypothesis (H1) states that there is a difference. Based on the results of the one-tail and two-tail tests, with a p-value $\geq$ 0.05, H0 is accepted, which means that there is no significant difference between the means of DPM and the modules.

4 Discussion

To get the accuracy, the module is compared with a digital pressure meter. Researchers have conducted research in making negative pressure wound therapy with PID control. The working system in this research, circuit, and software can work properly. The circuit consisting of ESP32 can work with a voltage of 5 V. In this module, researchers use the XGZP101DB1R sensor to monitor the pressure on the vacuum pump which functions as part of the vacuum pump control circuit. In addition, the system is equipped with a selenoid to close and open the vacuum flow. In terms of data processing, this module utilizes the HX710.h library to process the digital output of the XGZP101DB1R sensor. The data processing process involves a temperature query to obtain digital data from the XGZP101DB1R sensor. The data is then converted into mmHg pressure and stored in variables. In addition, for pressure control using the PID method, this program also uses the PID library. KP, KI, and KD constant settings are used as PID program parameters. The program works by calling the PID control command according to the predetermined settings. The PID control system is adjusted through an experiential approach or trial-and-error method. This condition arises because the parameters KP, KI, and KD are not independent. To achieve optimal control performance, experiments are required by

trying various combinations of P, I, and D until the desired values of KP, KI, and KD can be found.

Research conducted by Fikri Fahriansyah and team in 2020, with the title "A Low Cost Negative Pressure Wound Therapy Based on Arduino," utilizes MPXV4115VC6U and MPXV5050VC6T1 pressure sensors in the pressure range of 0 to −350 mmHg. The Arduino microcontroller is used to process data from the sensors, and the measurement results are displayed on a liquid crystal display (LCD) screen. However, the operation mode only uses continuous mode in the therapy process. Meanwhile, this research uses 3 modes of operation, namely continuous, intermittent, and dynamic, and uses the XGZP101DB1R pressure sensor to detect negative pressure and uses the PID method as control.

In addition, this research has several advantages including being able to read negative pressure with a pressure value of 0 to −150 mmHg, using a battery power source. So there is no need to look for an electric plug when used. The battery used can be recharged if the battery runs out, also equipped with a battery level indicator and an in-charge and full indicator. The shape is light, so it is easy to carry anywhere. However, this research also has shortcomings, namely the negative pressure reading value of −50 mmHg and below the vacuum pump is less stable, the LCD Nextion touchscreen button is still insensitive and sometimes the image is blank, the canister or tube where the liquid is less large, so that it can accommodate 1 L of liquid, the sound of the vacuum pump is still clearly audible when working.

The creation of a Negative Pressure Wound Therapy (NPWT) device with PID control and three operating modes of continuous, intermittent, and dynamic provides significant benefits to patients and the healthcare field. The PID control enables stable and precise pressure regulation, helping to accelerate the wound healing process by maximizing blood flow, reducing infection, and accelerating tissue growth. With three operating modes, medical personnel can customize therapy according to the patient's wound condition: continuous mode provides constant pressure for wounds that require steady suction, intermittent mode regulates pressure in time intervals to stimulate tissue response, and dynamic mode allows for more adaptive pressure changes as required by the wound condition. This increases therapeutic flexibility, patient comfort, and efficient use of medical resources, and optimizes faster and more effective healing outcomes.

5 Conclusion

Based on the results of research using the negative pressure wound therapy tool module using the PID system in continuous, intermittent, and dynamic modes, it can be concluded that: The creation of a PID control circuit on negative pressure wound therapy that can regulate the condition of the vacuum pump, when the vacuum pump is close to the setting pressure, the vacuum pump sucks more slowly where it is to maintain negative pressure conditions in accordance with what has been set. The NPWT tool is made using 3 modes, namely continuous, intermittent, and dynamic. The pressure sensor uses XGZP101DB1R to detect negative pressure. Equipped with a battery the tool is more portable and can be carried anywhere. After 5 measurements, the test results show that the Negative Pressure Wound Therapy (NPWT) prototype based on PID control can

achieve negative pressure according to the set point with the largest average error of 0%, 0%, and 1.13% and the lowest error of 0%, 0%, and 0.31% in continuous, intermittent, and dynamic modes, respectively. The highest error rate in pressure measurement using a Digital Pressure Meter (DPM) is 2.33%, and the lowest is 0.15% in all modes. The prototype was able to achieve set pressure at varying times, depending on the mode of operation, with an average time that was fast enough to achieve pressure stability. The pressure range tested was −50 to −150 mmHg, which corresponds to the needs of NPWT therapy.

The development of a more affordable and portable NPWT can provide an effective solution in the treatment of diabetic ulcers, which are a serious complication in diabetic patients. The use of PID control in this prototype ensures more precise pressure regulation, thus improving therapeutic effectiveness and patient comfort. The portable battery-equipped prototype allows for easy mobility, making it useful in both clinical and home environments. With minimal measurement error, it has the potential to be widely used in wound care with high accuracy. A further development that can be done is to use a more silent and smoother type of vacuum pump when the vacuum pump rises more to make it more stable. Adding a leak reading sensor to make it easier to control the hose connector. The tool is made smaller so that it can stick to the body so that patients can move for small activities such as walking.

References

Patel S, Srivastava S, Singh MR, Singh D (2019) Mechanistic insight into diabetic wounds: pathogenesis, molecular targets and treatment strategies to pace wound healing. https://doi.org/10.1016/j.biopha.2019.108615

Vijayakumar V, Samal SK, Mohanty S, Nayak SK (2019). Recent advancements in biopolymer and metal nanoparticle-based materials in diabetic wound healing management. https://doi.org/10.1016/j.ijbiomac.2018.10.120

Yassin AM, Mohamed M, Dirar M, Ahmed M, Elsir K, Alhadi A (2024). Modified intermittent low negative pressure wound therapy for complex injuries: a case series. https://doi.org/10.1016/j.ijscr.2024.109356

Griffin DR, Weaver WM, Scumpia PO, Di Carlo D, Segura T (2015) Accelerated wound healing by injectable microporous gel scaffolds assembled from annealed building blocks. Nat Mater 14:737–744. https://doi.org/10.1038/nmat4294

Grainger DW (2015). Wound healing: enzymatically crosslinked scaffolds. https://doi.org/10.1038/nmat4337

Liu D, Zhang H, Fontana F, Hirvonen JT, Santos HA (2017). Microfluidic-assisted fabrication of carriers for controlled drug delivery. https://doi.org/10.1039/c7lc00242d

Li J et al (2018). Nanocellulose-Based Antibacterial Materials. https://doi.org/10.1002/adhm.201800334

Lai WF, Shum HC (2015) Hypromellose-graft-chitosan and its polyelectrolyte complex as novel systems for sustained drug delivery. ACS Appl Mater Interfaces 7:10501–10510. https://doi.org/10.1021/acsami.5b01984

Chi J, Zhang X, Chen C, Shao C, Zhao Y, Wang Y (2020) Antibacterial and angiogenic chitosan microneedle array patch for promoting wound healing. Bioact Mater 5:253–259. https://doi.org/10.1016/j.bioactmat.2020.02.004

Mohandas A, Anisha BS, Chennazhi KP, Jayakumar R (2015) Chitosan-hyaluronic acid/VEGF loaded fibrin nanoparticles composite sponges for enhancing angiogenesis in wounds. Colloids Surf B Biointerfaces 127:105–113. https://doi.org/10.1016/j.colsurfb.2015.01.024

Anisha BS, Biswas R, Chennazhi KP, Jayakumar R (2013) Chitosan-hyaluronic acid/nano silver composite sponges for drug resistant bacteria infected diabetic wounds. Int J Biol Macromol 62:310–320. https://doi.org/10.1016/j.ijbiomac.2013.09.011

Yao S, Luo Y, Wang Y (2022). Engineered microneedles arrays for wound healing. https://doi.org/10.1016/j.engreg.2022.05.003

Othman Z, Cillero Pastor B, van Rijt S, Habibovic P (2018). Understanding interactions between biomaterials and biological systems using proteomics. https://doi.org/10.1016/j.biomaterials.2018.03.020

Cortell-Fuster C, Gaspar-Carreño M, Achau-Muñoz R, Delgado-Ruiz T, Hortelano-Otero A (2019) Negative pressure therapy with instillation for the treatment of infected wounds: recommendations of utilization based on evidence. Farm Hosp 43:6–12. https://doi.org/10.7399/fh.11011

Borgquist O (2013) Negative pressure wound therapy : therapy settings and biological effects in peripheral wounds. Lund University

Karatepe O et al (2011) Vacuum assisted closure improves the quality of life in patients with diabetic foot. Acta Chir Belg 111:298–302. https://doi.org/10.1080/00015458.2011.11680757

Uğurlar M, Sönmez MM, Armağan R, Eren OT (2017) Comparison of two different vacuum-assisted closure (VAC) treatments of multiple chronic diabetic foot wounds in the same extremity. Foot Ankle Surg 23:173–178. https://doi.org/10.1016/j.fas.2016.05.314

Chen L, Zhang S, Da J, Wu W, Ma F, Tang C, Li G, Zhong D, Liao B (2021) A systematic review and meta-analysis of efficacy and safety of negative pressure wound therapy in the treatment of diabetic foot ulcer. Ann Palliat Med 10:10830–10839. https://doi.org/10.21037/apm-21-2476

Farina JA, de Almeida CEF, Marques EGSC, Jorge JLG, Lima RVKS (2015) Letter to the editor: negative pressure wound therapy in grade IIIB tibial fractures: fewer infections and fewer flap procedures? https://doi.org/10.1007/s11999-015-4477-5

Almeida IR, Coltro PS, Gonçalves HOC, Westin AT, Almeida JB, Lima RVKS, Silva MF, Farina Junior JA (2021) The role of negative pressure wound therapy (NPWT) on the treatment of pyoderma gangrenosum: a systematic review and personal experience. https://doi.org/10.1111/wrr.12910

Zurovcik D, Slocum A, Mody G, Riviello R, Sheridan R (2011) Development of simplified negative pressure wound therapy device for low-resource settings. In: Proceedings—2011 IEEE global humanitarian technology conference, GHTC 2011, pp 91–97. https://doi.org/10.1109/GHTC.2011.23

Lee KN, Ben-Nakhi M, Park EJ, Hong JP (2015) Cyclic negative pressure wound therapy: an alternative mode to intermittent system. Int Wound J 12:686–692. https://doi.org/10.1111/iwj.12201

Niu H, Tian H, Chang C, Li D, Fan Y (2011) Development of a device for negative pressure wound therapy based on ARM9 embedded system. In: 2011 5th International conference on bioinformatics and biomedical engineering. IEEE, pp 1–4. https://doi.org/10.1109/icbbe.2011.5780307

Tan J-H, Tan T-S, Rafiq Abdul-Kadir M, Saad AZM, Teoh C-K, Javier J, Olmedo S, Lum K-Y, Leong K-M, Iskandar TA, Alang T, Malik SA (2017) Incorporating fuzzy logic into an adaptive negative pressure wound therapy device. J Telecommun, Electron Comput Eng 9(4)

Pramono FF, Luthfiyah S, Rahmawati T, Hasanah Ahniar N (2020) A low cost negative pressure wound therapy based on Arduino. Indonesian Journal of electronics, electromedical engineering, and medical informatics 2:68–73. https://doi.org/10.35882/ijeeemi.v2i2.2

Amanah AU (2022) Negative Pressure Wound Therapy Dilengkap 2 Mode

Davida A (2023) Development of simple control system for a negative pressure wound therapy device. Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI) 9:1080–1092. https://doi.org/10.26555/jiteki.v9i4.27152

Carroll CA, Moore BL, Ingram SC, Justin N (2018) Wound therapy system with instillation therapy and dynamic pressure control. https://patents.google.com/patent/US20200078224A1/en

Christopher Brian L, Justin Alexander L (2024) Fluid instillation apparatus for use with negative-pressure system incorporating wireless therapy monitoring, https://patents.google.com/patent/US20240033131A1/en?q=(fluid+instillation+apparatus)&oq=fluid+instillation+apparatus

Andreas A, Priyandoko G, Mukhsim M, Putra SA (2020) KENDALI KECEPATAN MOTOR POMPA AIR DC MENGGUNAKAN PID—CSA BERDASARKAN DEBIT AIR BERBASIS ARDUINO. JASEE Journal of Application and Science on Electrical Engineering 1:1–14. https://doi.org/10.31328/jasee.v1i01.3

Carroll CA, Antonio S (2018) Wound therapy system with instillation therapy and dynamic pressure control. https://patents.google.com/patent/US20200078224A1/en

Li Z (2023) Review of PID control design and tuning methods. In: Journal of physics: conference series. Institute of Physics. https://doi.org/10.1088/1742-6596/2649/1/012009

Design Simulation of a Voltage Balancing System in a Series Photovoltaic Array Based on an Isolated Bi-directional Converter with the Proportional-Integral Controller

Mukhammad Nur Hadi[1], Fahriza Balqis[1], Arya Kusumawardana[1(✉)], and Adi Izhar Che Ani[2]

[1] State University of Malang, Malang, East Java 65145, Indonesia
arya.kusumawardana.ft@um.ac.id
[2] Universiti Teknologi MARA, Shah Alam, Pulau Pinang, Malaysia

Abstract. This study explores the design and simulation of a voltage balancing system for series photovoltaic (PV) arrays using an isolated bi-directional converter integrated with a Proportional-Integral (PI) controller. Partial shading on PV panels can result in mismatched output, leading to long-term degradation and reduced performance. The proposed system employs a PI-controlled flyback converter to balance voltage across individual PV modules, mitigating the effects of partial shading. Through simulations in MATLAB/Simulink, the research determines optimal PI parameters with $k_p = 0.09$ and $k_i = 5$. The results indicate that the designed controller significantly improves voltage convergency, yielding a system response with an RMSE 0.1495 better than a conventional controller. However, the designed system has a narrow voltage regulation workspace to generate MPP. Overall, the proposed system shows promising potential in enhancing PV array efficiency under varying shading conditions.

Keywords: Photovoltaic · Bi-Directional Flyback Converter · PI Control · Partial Shading

1 Introduction

Partial shading significantly affects photovoltaic (PV) system performance and lifespan, substantially reducing power output and efficiency. Shading 25%, 50%, and 75% of a module's surface cuts power by 33.7%, 45.1%, and 92.6%, respectively [1]. A 10% shading increase leads to a 12.41 W power drop and a 2.3% efficiency decrease [2]. The shading impact factor varies with irradiation, being 1.25 at 600 W/m^2 with 25% shading and 0.86 with 75% shading [2]. Water droplets on PV panels can enhance output by at least 5.6% due to cooling [1]. Partial shading also causes long-term module degradation [1] and hot-spot phenomena, reducing PV generator lifespan [3]. Researchers recommend efficient PV designs and advanced maximum power point tracking (MPPT) techniques to mitigate these issues and improve power yield under partial shading [3].

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 378–392, 2026.
https://doi.org/10.1007/978-981-95-9233-3_27

Differential power processing (DPP) enhances photovoltaic (PV) system efficiency and power output under mismatch conditions caused by partial shading, electrical disparities, or varying irradiance levels [4–6]. DPP converters maximize power extraction from each submodule, boosting energy harvesting and system efficiency, particularly under uneven lighting [7]. Various configurations, control strategies [8], and advanced algorithms [9–11] optimize performance, lower costs, improve reliability, and support standardized modular designs [12]. DPP is crucial for addressing mismatch challenges in PV systems, enabling efficient power extraction from individual submodules and minimizing conversion losses, ultimately enhancing the performance and energy yield of photovoltaic installations.

DPP in segmented photovoltaic (PV) systems optimizes power generation and efficiency under mismatched conditions. The segmented DPP structure, using bidirectional DPP flyback converters, suits large-scale systems, reducing installation complexity and high voltage ratings [13]. However, other topologies, such as the unit DPP structure combining bidirectional flyback and boost converters, address scalability and ensure maximum power point operation [7]. A new boost-flyback converter topology for feed-forward DPP systems also maintains voltage under partial shading [14]. The optimal topology depends on specific requirements. The segmented DPP approach achieves up to 96.4% efficiency in even lighting and 92.7% in severe partial shading [13], while unit DPP improves efficiency by 9–10% over series string under uneven lighting [15]. An equally important part of the topology is the control strategy.

Control strategies are vital for the development and efficiency of differential power processing (DPP) systems in photovoltaic (PV) applications. Numerous studies discuss control methods to enhance DPP performance. A voltage inner loop and power outer loop control method are suggested for tracking the maximum power point of PV panels while only converting differential power between mismatched modules [16]., reducing the coupling effects in modular DPP systems. Another method, a self-adjusted time-sharing maximum power point tracking (MPPT) control strategy, is designed for sub-module DPP, enabling precise and rapid tracking of all sub-module maximum power points with a single MPPT controller [11], streamlining system integration and lowering costs and conversion losses. Some strategies aim to minimize the power processed by DPP converters for higher efficiency. For example, a minimum-power tracking (MPT) algorithm is proposed for two-stage PV-PV DPP structures, where each PV executes MPPT control while central control employs MPT [10]. Another approach introduces a low-complexity power balancing point-based optimization algorithm to cut system costs and size, boost efficiency, and promote standardized modular DPP converter designs [12]. Thus, control strategies in DPP systems are crucial for maximizing power output, minimizing losses, and enhancing overall system efficiency, addressing challenges such as mismatch conditions, scalability, and cost-effectiveness in PV systems using DPP technology.

Control strategies for differential power processing (DPP) in photovoltaic (PV) systems have improved efficiency and power output. The isolated-port DPP architecture control strategy drives the duty-cycle of differential submodule integrated converters (subMICs) proportionally to the voltage difference between the submodule and the isolated-port, requiring no additional sensing, processing, or communication [17]. However, contradictions exist in control strategies, such as multi-loop control using proportional integral (PI) controllers for load voltage regulation [18] and voltage inner loop and power outer loop control for modular DPP [19, 20]. Recent strategies focus on achieving maximum power point tracking (MPPT) while minimizing power losses, with approaches like the self-adjusted time-sharing MPPT control strategy [11] and the minimum-power tracking (MPT) algorithm for two-stage PV-PV DPP structure [10] aiming to improve efficiency under various conditions, including partial shading.

The DPP's performance is affected by both control strategy and topology selection. Research [21] has shown success in reducing submodule mismatches to below 25% using a linear controller and bidirectional flyback converter (BFC). Efficiency gains under shading conditions have been realized through nonlinear controllers [22]. With the same topology, [23] prioritized achieving MPP for each submodule, resulting in better performance than the DPP-less system. Additionally, the open-loop voltage equalization approach with constant duty ratio in [24, 25] has effectively increased power generation from PV modules under shading. Although prior studies have demonstrated enhanced PV output through various control methods, there is still a lack of comprehensive analysis on voltage imbalance issues between submodules.

Based on the studies that have been done, the PI controller is one of the simple controllers and is able to be used in modular or segmented topology. Moreover, it is simple, effective, and suitable for controlling bidirectional flyback converters, offering fast and accurate power flow regulation while addressing steady-state errors. In this study, a PI controller adjusts the flyback converter's duty cycle based on the differential power between modules, ensuring balanced power flow and optimal PV array efficiency. The controller design considers the flyback converter's dynamic behavior and PV module characteristics. Simulations demonstrate the PI-controlled system's effectiveness in mitigating partial shading effects and enhancing overall PV array efficiency.

This research primarily contributes by designing and simulating a modular DPP system based on a bidirectional flyback converter integrated with a PI controller, addressing a gap in the literature. System performance was also analyzed based on the ability to reduce voltage imbalance between sub-modules. Additionally, the article analyzes the dynamic behavior of the PI-controlled flyback converter in a DPP system, offering insights into stability and performance under partial shading. The designed controller is also compared with other controllers to determine its performance. The voltage balancing mechanism is further discussed in the next subsection

2 Voltage Balancing Mechanism

The proposed DPP, depicted in Fig. 1, comprises a series circuit, also known as PV-to-Isolated Port (PV-IP) or PV-to-Virtual Bus [26], featuring bidirectional circuits. The primary section of each circuit connects to a PV sub-module element, while the secondary

section links to the capacitor section, serving as an isolated port. Mathematically, the isolated port must maintain a balance between incoming and outgoing power as seen in Eq. (1).

$$\sum_{k=1}^{n} V_{PV,k} I_{DPP,k} = 0 \tag{1}$$

where $I_{DPP,\,k}$ and $V_{PV,\,k}$ are the current through the DPP and the PV voltage, respectively, for $k = 1, 2, \ldots, n$ indicate the number of elements for both the PV sub-module and the DPP. Thus, the string current can be defined as in Eq. (2).

$$I_{string} = I_{PV,k} - I_{DPP,k} \tag{2}$$

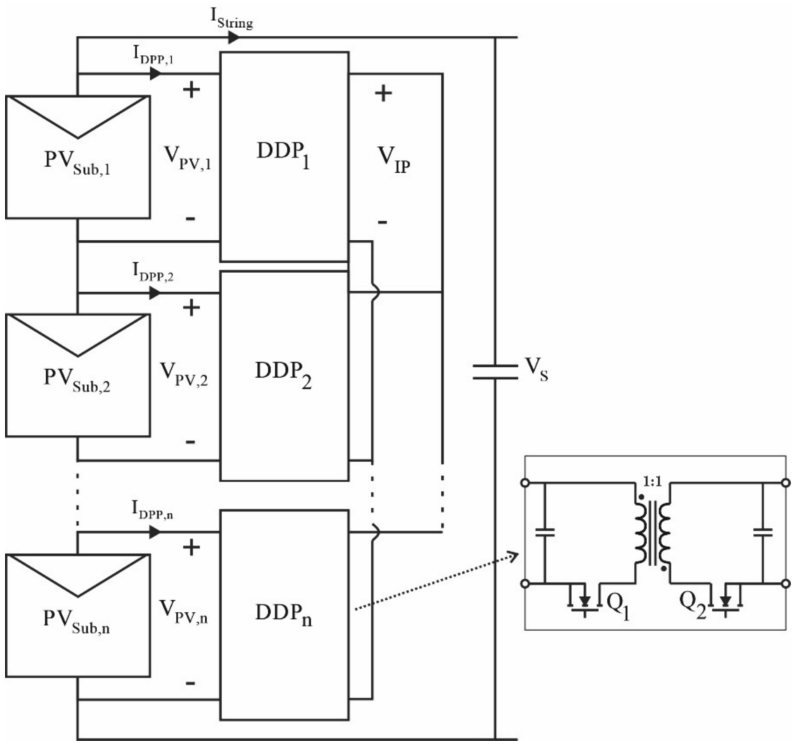


Fig. 1. Simplified PV-to-isolated port architecture

To function correctly, the DPP converter must utilize a bidirectional isolated topology. This design offers the benefit of allowing each PV element's power to be directly exchanged with a separate bus that operates at a lower voltage than the main bus voltage, thereby reducing the voltage ratings and costs of components. Nevertheless, to ensure the isolated bus maintains a stable voltage, the power flowing into and out of the bus

must be balanced, as indicated in Eq. (1). Consequently, achieving proper MPPT may not be possible at certain string currents, which represents a compromise inherent in selecting this architectural approach [27].

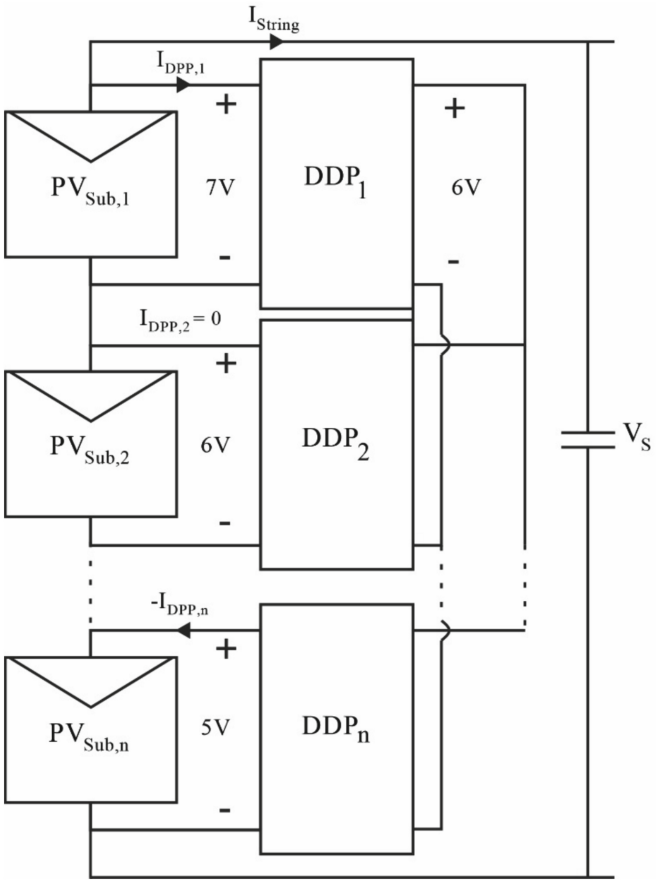


Fig. 2. Operation principle of PV-IP

In this case, voltage balancing is prioritized overachieving MPP. It is thus assumed that the designed device does not accommodate the control of strategies to achieve MPP. In simple terms, the voltage balancing mechanism is illustrated in Fig. 2. Each DPP is an independent unit that does not exchange data. PV voltage balancing in this case can also be interpreted as balancing the capacitor voltage in the primary section. In this case, the communication less condition on each DPP must be fulfilled while balancing the voltage of each sub-module. The mechanism that can be done is to balance the primary voltage with the voltage on the isolated port. Power transfer between the primary section and the IP section can be done by adjusting the duty ratio of Q1 and Q2. Both MOSFETs work in complement. If there is a PV with more voltage than the IP, then the DPP will flow current

from the PV to the IP and vice versa. To ensure the fulfilment of the balanced condition, the design of the model and controller is further explained in the next sub-section.

3 Model Design

3.1 Design of Simulation Model

To determine the performance of the proposed system, this research utilizes MATLAB/Simulink to test numerically. The computer specifications used in the test are using AMD Ryzen 7 8845 HS 3.8 Ghz Processor, 16GB RAM and SSD with Sequential Read 4000 MB/S, Sequential Write 3000 MB/S.

The simulation model built is shown in Fig. 3. There are two PVs and two DPPs in the test. The PV strings are connected to the DC string converter device and the load to generate MPP (not discussed in this case). One of the important mechanisms to be done before testing the system is the design of the DPP. As shown in Fig. 1, there is a bidirectional flyback converter as the main circuit of the DPP. The parameters of the DPP are shown in Table 1.

Table 1. Key parameter of simulation design

Parameter	Symbol	Value	Unit
Nominal IP voltage	V_{IP}	6	V
Switching frequency	f_{sw}	120	kHz
Maximum current flow through DPP	$I_{DPP,\,n}$	2	A
Ripple IP voltage	ΔV_{IP}	5	%
Ripple magnetizing inductance current	Δi_{L_m}	10	%
Nominal duty cycle	D	50	%

Based on the parameters shown in Table 1 and the literature [28], the component parameters are obtained as shown in Table 2. Each parameter is used as a reference to give a value to each passive component including transformers in the simulation model. The PV cell parameters are shown in Table 3. These parameters are single cell parameters, in this case one sub-module uses 10 cells assembled in series. So that the I-V and P-V curves under conditions exposed to 1000 W/m^2 radiation are shown in Fig. 4.

The curve shown in Fig. 4 is an idealized curve under non-shaded conditions. In this case, the shading condition is tested for one sub-module rather than one cell. So, the voltage value under ideal conditions and reaching MPP is around 6 V with a current close to 10 A. In the next section, we will discuss the controller design.

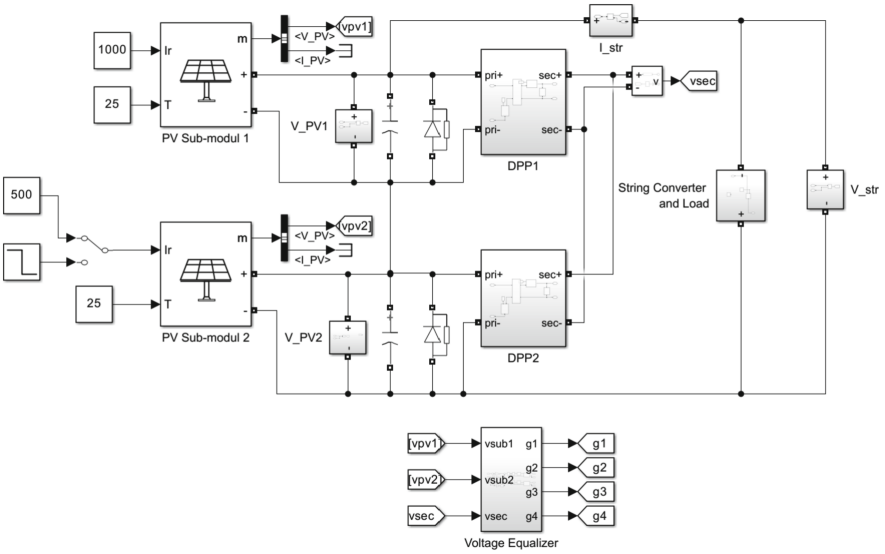


Fig. 3. Design of simulation model.

3.2 Controller Design

Figure 5 is the controller model used in this study, where the control system used is a PI control system. In general, the formula for PI controller is shown in Eq. (3).

$$u_{pi,k} = \varepsilon_k \left(k_{p,k} + \frac{1}{s} k_{i,k} \right) \tag{3}$$

where ε is the difference between the voltage of sub-module k and the voltage of IP, $\varepsilon = V_{PV,\,k} - V_{IP}$. If D_k is the duty ratio for each DPP, then D_k can be written as in Eq. (4).

$$D_k = \left| \text{sat}\left(u_{pi}\right) \right| \tag{4}$$

where the sat(·) is a saturation function as shown in Eq. (5).

$$\text{sat}\left(u_{pi}\right) = \begin{cases} u_{pi} > 0.95 & u_{pi} = 0.95 \\ -0.95 \le u_{pi} \le 0.95 & u_{pi} = u_{pi} \\ u_{pi} < -0.95 & u_{pi} = -0.95 \end{cases} \tag{5}$$

The function in Eq. (5) is needed to provide a limit on the duty ratio value generated by the PI controller. This is because the converter has a duty ratio value limitation of less not more than 0.95.

The utilization of PI controller leaves issues related to the determination of parameters k_p and k_i. Since the proposed DPP system is complex, a linear system approach is not suitable in this case. An empirical approach is used to determine both parameters as opposed to analytical nonlinearity. The tuning mechanism can be done as follows:

- Set the k_p value from 0 to a certain value and choose the k_p value with the best performance with $k_i = 0$.

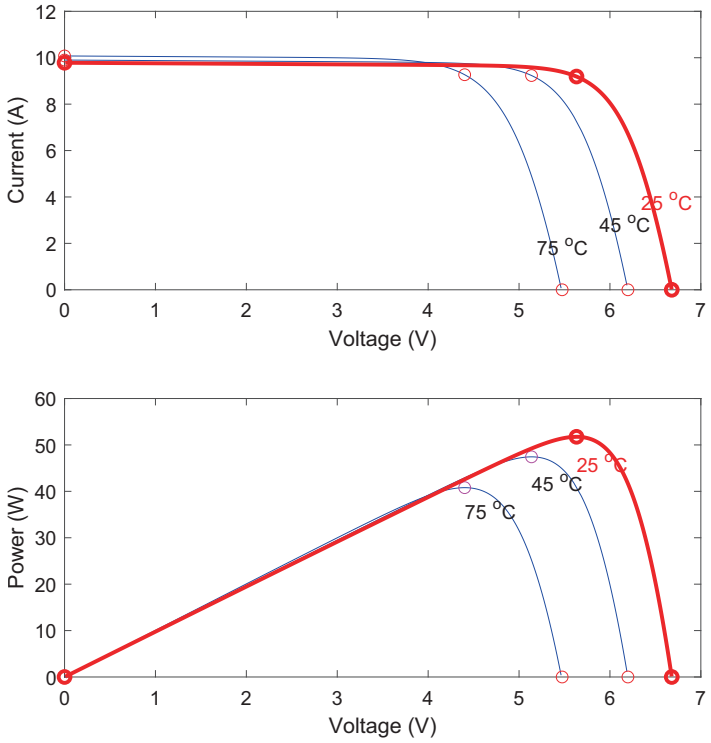


Fig. 4. I-V dan P-V curves of PV sub-module under test in 1000 W/m^2

Table 2. The converter parameters value.

Parameters	Symbol	Value	Unit
Magnetizing inductance	L_m	31.25	H
Number of turn primary winding	n_1	30	uH
Number of turn secondary winding	n_2	30	uH
Turn ratio	N	1	–
PV capacitor	$C_{PV,\,n}$	220	uF
IP capacitor	C_{IP}	100	uF

- Using the optimal k_p value and setting the k_i value from 0 to a certain value and selecting the value with the best performance.
- System performance is determined by the lowest RMSE value of ε.
- The PI controller tuning process is performed by providing a constant and different radiation value between sub-modules 1 and 2.

Table 3. The PV cell parameters value.

Parameters	Symbol	Value	Unit
Maximum power	P_{max}	5.174	W
Open circuit voltage	V_{oc}	0.668	V
Short circuit current	I_{sc}	9.7831	A
Voltage @MPP	V_{mp}	0.5632	V
Current @MPP	I_{mp}	9.1892	A
Temperature coefficient of V_{oc}	α	−0.36099	%/°C
Temperature coefficient of I_{sc}	β	0.06	%/°C

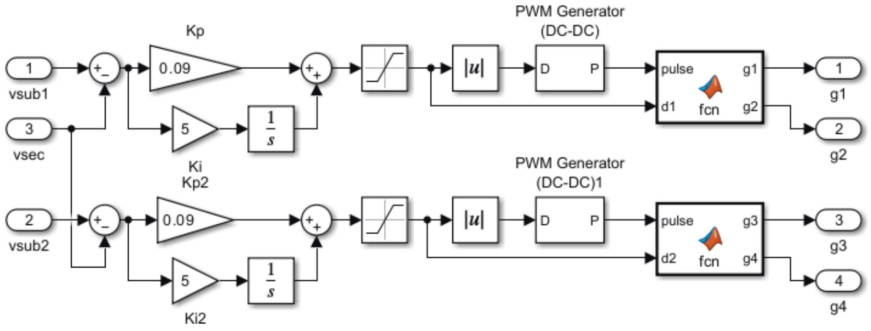


Fig. 5. Controller design.

Proportional Gain Tuning

Testing the k_p value is done by fine tuning, where testing is done by giving the k_p value ranging from 0.01 to 0.2 and with a Ki value of 0 (zero) and with a test time of 0.5 s. The k_p value used was chosen to get the bandwidth value between the two inputs or the error value with the lowest value, as for if the k_p value is negative then the resulting system response becomes isolated as shown in Fig. 6 so that the use of the k_p value must be positive.

Based on the trial results shown in Fig. 7a, the lowest RMSE value is obtained with a k_p value of 0.1 with the three lowest values being 0.09, 0.1, and 0.11. The three values are selected to determine the k_i value.

Integral Gain Tuning

Testing the k_i value is done by fine tuning, where the test is carried out by giving a k_i value of 0.1 to 5. The test was carried out three times with three k_p values that had been selected in the previous. Based on Fig. 7b, the change in the value of k_i for each k_p value that has been obtained is not much different for the RMSE value. Interestingly, the greater the value of k_i, the smaller the RMSE value. The selected k_i value is 5 because there is no significant change above the value of 4.

This value will be used in the next testing process to find out the difference in system response that will be generated in testing the system response by manipulating the load.

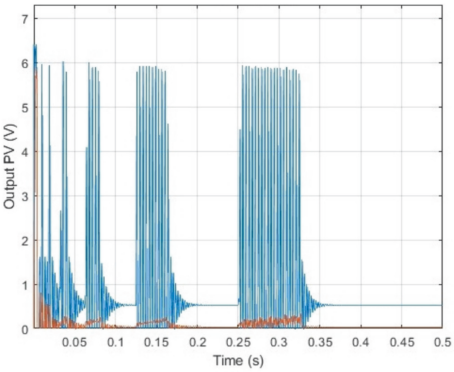


Fig. 6. System response with negative k_p.

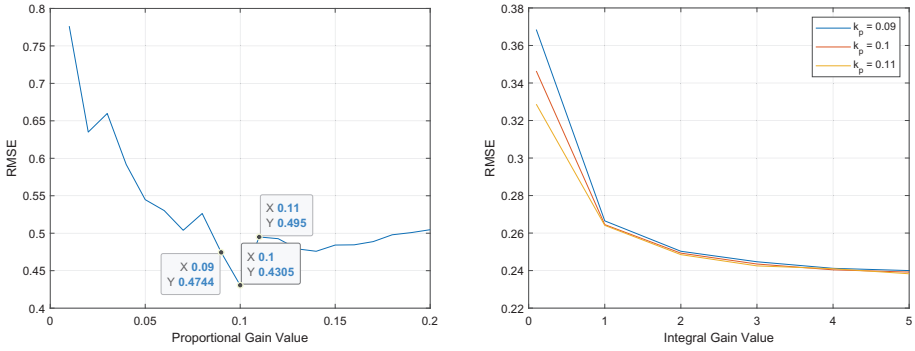
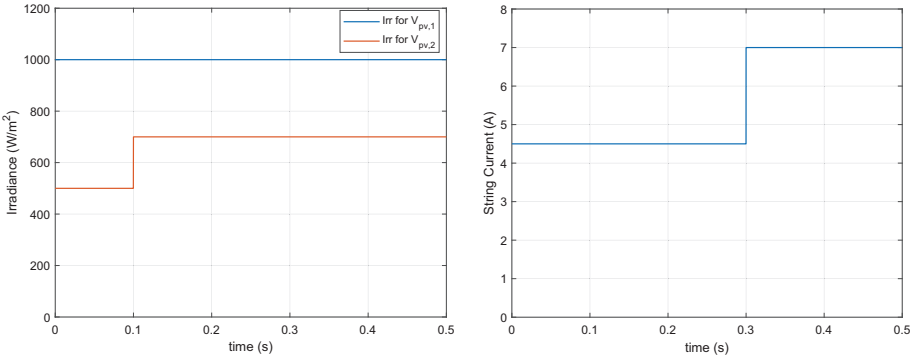


Fig. 7. RMSE value; (**a**) with varied proportional gain; (**b**) with varied integral gain

4 System Performance Result

System performance testing is a crucial part that must be done to know the performance of the proposed system. In this case, the testing is done with the same ecosystem. There are two test discussions, firstly related to the system response test and secondly related to the P-V curve generated by the system as shown in Fig. 9. The test was conducted with time sampling, t_s, with a value of 1e-6 s. The solver type is fixed-step with ode8 (Dormand-Prince).

The voltage balancing system that is built must be able to maintain balance when there are changes in irradiation and strong currents. This is in accordance with the conditions in the real environment that the DPP device will be integrated with the MPPT device where to achieve MPP the system will produce a dynamic string current. In this work, the dynamic irradiance and string current changes are shown in Fig. 8.



(a) (b)

Fig. 8. Dynamics behavior; (**a**) Irradiance value of PV1 and PV2; (**b**) PV string current

The proposed controller's impact on PV1 and PV2 voltage output is examined. Both Vpv1 and Vpv2 reach a steady state of 6 V, demonstrating a smooth response with minimal voltage disparity. The system rapidly equalizes the voltages, exhibiting brief initial oscillations before quickly stabilizing. In contrast, the proportional controller shows a slower convergence rate than the proposed method, with more noticeable differences between $V_{PV,\,1}$ and $V_{PV,\,2}$ over time. Although the system eventually stabilizes, the voltage discrepancy remains more evident compared to the proposed controller's results. A significant shift in string current at t = 0.3 s causes observable voltage fluctuations.

Initially, the discrepancy between $V_{PV,\,1}$ and $V_{PV,\,2}$ is somewhat elevated but swiftly diminishes to almost zero, indicating the controller's efficiency in reducing the voltage difference. The system employing the proportional controller exhibits a higher normalized square error compared to the suggested controller, and this error remains evident throughout the simulation period. The proportional controller fails to minimize the error as efficiently as the proposed controller does. In more detail, the proposed controller produces a system response with an RMSE of 0.1495 while the proportional controller produces 0.2574. Thus, it is known that the proposed controller produces a better response due to the smaller RMSE value.

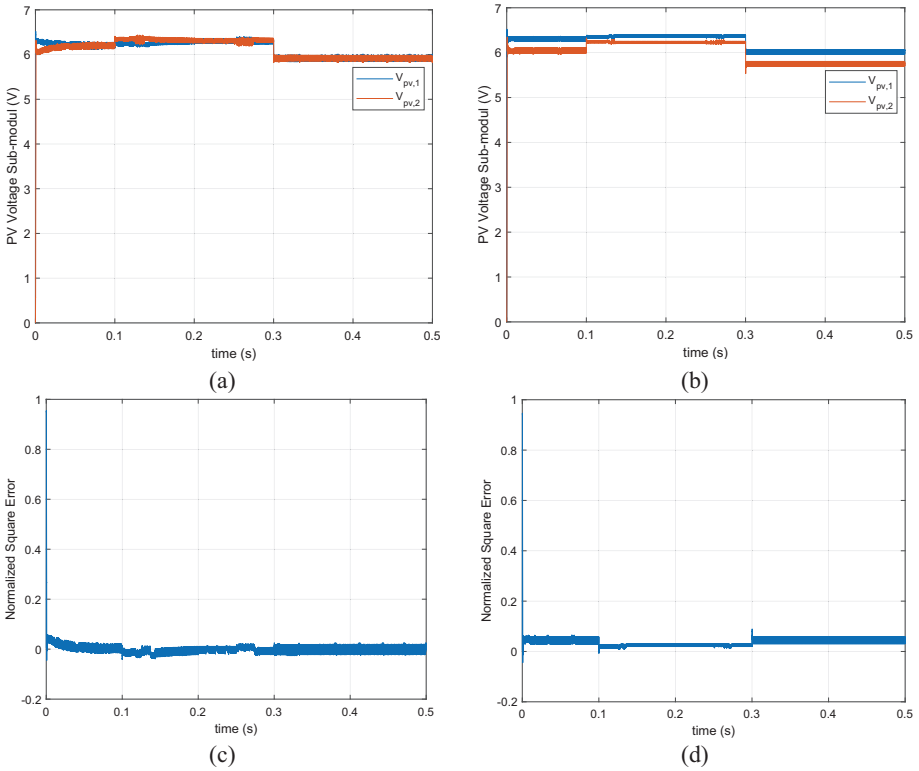


Fig. 9. System responses; (**a**) system response with proposed controller; (**b**) system response with proportional gain control; (**c**) normalized ε system with proposed controller; (**d**) normalized ε system with proportional gain controller.

In Fig. 10, it is defined that P_1, P_2, and P_3 are the power curves of the proposed system, a system with a proportional controller and a system without DPP, respectively. The maximum power generated by the proposed system is greater than the other two with a value of 74.57 W. This value is 19.67 W greater than the power without DPP. While the power generated by the second system is 12.62 W greater. The significant difference in the proposed system is not the only positive outcome of this work. In addition, the proposed system was also able to eliminate the condition of the curve with local maximum. Moreover, the second system produces abnormal P-V curves at voltages between 4 V and 8 V. This will greatly impact the operation of the string converter to reach MPP. Such unusual conditions can trap the MPPT algorithm in conditions that are not global maximum. Although the P-V curve generated by the proposed system is better than its competitors, the condition of no data at a voltage of 2 V to 10 V will make MPPT operation on this system very limited. If the MPPT algorithm selection is not appropriate, it can cause system instability.

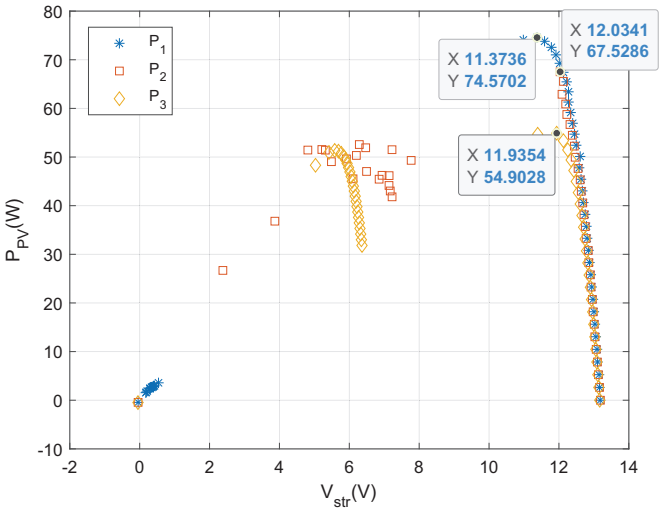


Fig. 10. Comparison of P-V curve

5 Conclusion

This study effectively develops and tests a voltage equalization system for photovoltaic (PV) arrays connected in series that are impacted by partial shading. The system employs an isolated bi-directional flyback converter with a Proportional-Integral (PI) controller. Partial shading poses substantial challenges to PV systems, including voltage disparities and power inconsistencies, which can result in gradual deterioration, localized overheating, and potentially irreversible damage to modules. The implemented system addresses these issues by maintaining balanced voltage across each PV module, thus enhancing the system's reliability and performance.

The MATLAB/Simulink experiments reveal that the PI controller effectively stabilizes the output voltage when optimal parameters are set to $k_p = 0.09$ and $k_i = 5$. Findings indicate that the PI controller, in comparison to proportional-only control, substantially minimizes voltage disparities between modules and provides more fluid dynamic performance. The proposed system achieves a lower RMSE value of 0.1495 than the competing system.

The research acknowledges a significant drawback in the proposed system: the designed system has a narrow voltage regulation workspace to generate MPP. Additionally, although the system design emphasizes voltage balancing, it may not consistently achieve Maximum Power Point Tracking (MPPT) when integrated with a string converter. Subsequent research could investigate the incorporation of sophisticated control mechanisms, such as adaptive or nonlinear controllers, to enhance the system's resilience. Moreover, conducting tests in authentic environmental settings and further refining MPPT algorithms will boost the practical viability of the suggested system.

References

1. Mustafa RJ, Gomaa MR, Al-Dhaifallah M, Rezk H (2020) Environmental impacts on the performance of solar photovoltaic systems. Sustainability 12:608. https://doi.org/10.3390/su12020608
2. Al Mamun MA, Hasanuzzaman M, Selvaraj J (2017) Experimental investigation of the effect of partial shading on photovoltaic performance. IET Renew Power Gener 11:912–921. https://doi.org/10.1049/iet-rpg.2016.0902
3. Belhaouas N et al (2021) A new approach of PV system structure to enhance performance of PV generator under partial shading effect. J Clean Prod 317:128349. https://doi.org/10.1016/j.jclepro.2021.128349
4. Amaral da Luz CM, Ribeiro ER, Tofoli FL (2021) Differential power processing applied to a photovoltaic string with different modules. In: 2021 Brazilian power electronics conference (COBEP). IEEE, pp 1–6. https://doi.org/10.1109/COBEP53665.2021.9684040
5. Amaral M, da Luz C, Moreira Vicente E, Lessa Tofoli F, E RR (2023) Differential power processing architecture to increase energy harvesting of photovoltaic systems under permanent mismatch. Sol Energy 263:111940. https://doi.org/10.1016/j.solener.2023.111940
6. Mahmoud Y (2022) Model-based approach for differential power processing (DPP) converters. In: IECON 2022—48th annual conference of the IEEE industrial electronics society. IEEE, pp 1–7. https://doi.org/10.1109/IECON49645.2022.9968493
7. Jeong H, Cho H-T, Kim T, Liu Y-C, Kim KA (2018) A scalable unit differential power processing system design for photovoltaic applications. In: 2018 IEEE 19th workshop on control and modeling for power electronics (COMPEL). IEEE, pp 1–8. https://doi.org/10.1109/COMPEL.2018.8460157
8. Badawy M, Sozer Y (2017) Differential power processing of photovoltaic systems for high energy capture and reduced cost. In: 2017 IEEE energy conversion congress and exposition (ECCE). IEEE, pp 475–481. https://doi.org/10.1109/ECCE.2017.8095821
9. Chu G, Wen H, Yang Y, Wang Y (2020) Elimination of photovoltaic mismatching with improved submodule differential power processing. IEEE Trans Ind Electron 67:2822–2833. https://doi.org/10.1109/TIE.2019.2908612
10. Ye Z, Wen H, Chu G, Li X (2017) Minimum-power-tracking for PV-PV differential power processing systems. In: 2017 IEEE 6th international conference on renewable energy research and applications (ICRERA). IEEE, pp 696–700. https://doi.org/10.1109/ICRERA.2017.8191150
11. Zhu T et al (2016) Sub-module differential power processing photovoltaic optimizer using self-adjusted time-sharing maximum power point tracking control strategy. In: 2016 18th European conference on power electronics and applications (EPE'16 ECCE Europe). IEEE, pp 1–6. https://doi.org/10.1109/EPE.2016.7695303
12. Chu G, Wen H, Hu Y, Jiang L, Yang Y, Wang Y (2020) Low-complexity power balancing point-based optimization for photovoltaic differential power processing. IEEE Trans Power Electron 35:10306–10322. https://doi.org/10.1109/TPEL.2020.2977329
13. Jeong H, Park S, Jung J-H, Kim T, Kim A-R, Kim KA (2021) Segmented differential power processing converter unit and control algorithm for photovoltaic systems. IEEE Trans Power Electron 36:7797–7809. https://doi.org/10.1109/TPEL.2020.3044417
14. Park J-H, Kim K-T (2017) Multi-output differential power processing system using boost-flyback converter for voltage balancing. In: 2017 international conference on recent advances in signal processing, telecommunications & computing (SigTelCom). IEEE, pp 139–142. https://doi.org/10.1109/SIGTELCOM.2017.7849811
15. Jeong H, Cho H-T, Kim T, Liu Y-C, Kim KA (2018) A scalable unit differential power processing system design for photovoltaic applications. In: 2018 IEEE 19th workshop on

control and modeling for power electronics (COMPEL). IEEE, pp 1–8. https://doi.org/10.1109/COMPEL.2018.8460157
16. Liu C, Zheng Y, Li D, Lehman B (2018) Distributed MPPT for modular differential power processing in scalable photovoltaic system. In: 2018 IEEE applied power electronics conference and exposition (APEC). IEEE, pp 1098–1103. https://doi.org/10.1109/APEC.2018.8341153
17. Levron Y, Clement DR, Choi B, Olalla C, Maksimovic D (2014) Control of submodule integrated converters in the isolated-port differential power-processing photovoltaic architecture. IEEE J Emerg Sel Top Power Electron 2:821–832. https://doi.org/10.1109/JESTPE.2014.2326972
18. Meshram K, Ghosh S, Jain S (2022) Optimal power management of off-grid solar photovoltaic system. In: 2022 Second international conference on power, control and computing technologies (ICPC2T). IEEE, pp 1–6. https://doi.org/10.1109/ICPC2T53885.2022.9776970
19. Liu C, Zheng Y, Lehman B (2022) PV panel to PV panel transfer method for modular differential power processing. IEEE Trans Power Electron 37:4764–4778. https://doi.org/10.1109/TPEL.2021.3123450
20. Liu, C.: Modular differential power processing in solar system, (2019). https://doi.org/10.17760/D20325265.
21. Olalla C, Clement D, Rodriguez M, Maksimovic D (2013) Architectures and control of submodule integrated dc-dc converters for photovoltaic applications. IEEE Trans Power Electron 28:2980–2997. https://doi.org/10.1109/TPEL.2012.2219073
22. Levron Y, Clement D, Maksimovic D, Olalla C Nonlinear control design for the photovoltaic isolated-port architecture with submodule integrated converters
23. Bell R, Pilawa-Podgurski RCN (2015) Decoupled and distributed maximum power point tracking of series-connected photovoltaic submodules using differential power processing. IEEE J Emerg Sel Top Power Electron 3:881–891. https://doi.org/10.1109/JESTPE.2015.2475607
24. Chu G, Wen H, Jiang L, Hu Y, Li X (2017) Bidirectional flyback based isolated-port submodule differential power processing optimizer for photovoltaic applications. Sol Energy 158:929–940. https://doi.org/10.1016/j.solener.2017.10.053
25. Dong P, Wen H, Chu G, Yang Y, Wang Y (2019) Power rating analysis and protection for photovoltaic-isolated port based differential power processing systems. Sol Energy 193:458–472. https://doi.org/10.1016/j.solener.2019.09.092
26. Shenoy PS, Kim KA, Johnson BB, Krein PT (2013) Differential power processing for increased energy production and reliability of photovoltaic systems. IEEE Trans Power Electron 28:2968–2979. https://doi.org/10.1109/TPEL.2012.2211082
27. Jeong H, Lee H, Liu Y-C, Kim KA (2019) Review of differential power processing converter techniques for photovoltaic applications. IEEE Trans Energy Convers 34:351–360. https://doi.org/10.1109/TEC.2018.2876176
28. Erickson RW, Maksimović D (2020) Fundamentals of power electronics. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-43881-4

Is the Grid Ready? The Impact of Growing Electricity Demand on Low-Voltage Distribution

Abdul Halim Ikram Mohamed(✉), Ruzlaini Ghoni, and Mohd Tarmizi Ibrahim

Department of Electrical and Automation, Faculty of Engineering Technology, University College TATI, Kemaman, Terengganu, Malaysia
24D01001@siswa.uctati.edu.my

Abstract. Low-voltage distribution networks are crucial for delivering electricity from generation points to end users, ensuring a reliable and efficient power supply. These networks use various configurations, including radial, ring, and spur circuits, each with distinct implications for reliability, voltage quality, and network losses. Radial and spur circuits are favored for their simplicity and cost-effectiveness but are prone to single point failures due to the absence of alternative pathways. In contrast, ring circuits enhance reliability by allowing fault isolation and restoration through alternative routes, though they come with increased complexity and cost. The selection of a network configuration depends on multiple factors, such as the volume of electricity transported, network design, and required reliability levels. Key components of these networks include distribution transformers, feeders, distributors, switchgear, and protection devices, all of which work together to maintain voltage stability, minimize losses, and ensure safety. Effective organization and management of these components are essential for sustaining network performance and addressing operational challenges. In conclusion, the configuration of low-voltage distribution networks and its key components significantly impacts reliability, voltage quality, and network losses. Therefore, it is essential to consider these factors in order to develop an efficient and reliable low-voltage distribution networks for consumers. This study aims to provide a better understanding regarding low-voltage distribution networks based on existing information for future reference and practical applications.

Keywords: Low-Voltage Distribution Network · Load Demand · Load Profile

1 Introduction

Low-voltage distribution networks are a final link in the power system whereby it receives the medium voltage distribution network and step-down to the consumer's voltage. Delivering a reliable power supply in this network is very difficult as it is prone to faults and disturbances. Despite their importance, low-voltage distribution networks are often overlooked their significance in supporting economic growth [1]. A reliable low-voltage distribution is crucial for residential, commercial, and industrial activities. For example, in the businesses industry, it cannot operate efficiently that directly impact productivity and economic growth. Poor power quality causes voltage to be unstable where it can damage the equipment, loss of productivity and lead to economic losses [2].

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 393–401, 2026.
https://doi.org/10.1007/978-981-95-9233-3_28

Due to rapid technologies development, various sophisticated equipment such as electric vehicles, medical equipment, smart homes and industrial processes relies on electrical power supply. Therefore, low-voltage distribution networks must be designed and modernized to accommodate these demands. Upgrades and modernization of low-voltage distribution networks can facilitate the integration of renewable energy sources, energy storage and advanced grid management systems. However, the surge in electricity demand due to population growth, economic expansion and numerous industries consumed electricity are exerting a massive pressure on low-voltage distribution networks. The increased load is leading to overloading, voltage drops and equipment failures that will compromise network reliability and performance. Moreover, the challenges of network capacity expansion in urban and densely populated areas are making the problem even worst.

Apart from that, the increasing penetration of distributed generation (DG) sources, particularly solar and wind energy in low-voltage networks presents both opportunities and challenges. Even though these renewable sources contribute to a better sustainable energy mix, yet their intermittent nature poses significant issues for network stability and voltage control. The output power produced by solar and wind energy can cause voltage and frequency to be unstable in the grid. This problem of course can cause adverse effect in network loading, hence complicating the management of power quality and increasing the risk of over-loads on transmission lines. Therefore, an advanced control systems that can maintain the stability of the power grid are required [3]. The rapid growth in electric vehicle (EV) adoption has a significant increase in electricity demand. Such growth requires a robust network of charging stations for consumers. This is because there are many factors needed to consider for integrating EV charging infrastructure into existing low-voltage networks as they tend to have a potential peak load increases and voltage fluctuations that can strain the power grid [4].

Low-voltage distribution networks are facing significant challenges due to aging infrastructure with some component dating back to the mid-20^{th} century [5]. Aging infrastructure causes frequent power outages and affecting the overall reliability of the network. Older infrastructure can result in voltage fluctuations, reduced power quality and increased energy losses.

1.1 Research Question and Significance

Understanding how electricity demand directly influences the reliability, efficiency and sustainability of the power grid in low-voltage networks is crucial. With the increasing integration of low carbon technologies such as photovoltaic systems and electric vehicles, low-voltage distribution networks face new challenges. For example, without forecasting an accurate demand, it is difficult for the utilities to plan and design networks that meet future electricity need. Therefore, addressing this research question can help utilities to overcome these problems by providing insights into demand patterns and enabling the deployment of demand response programs such as load shifting, which can alleviate network constraints and reduce the need for traditional network reinforcements [6]. This review aims to fill the knowledge gap by providing a structured analysis of existing data on low-voltage networks, which is currently scattered and not systematically compared,

thereby offering representative values for grid parameters that can inform future research and practical applications.

2 Understanding Electricity Demand

Electricity demand is referred as the amount of electricity required by consumers to power either their household, commercial buildings or industries without compromising the generation, transmission and distribution network reliability [7]. Typically, it is measured in megawatts (MW). In contrast, electricity consumption is referred as the amount of electricity used over a certain period. It is typically measured in kilowatt-hours (kWh). There are differences between load demand and load consumption. Demand is the measure of the rate at which power is needed at a specific time, indicating the immediate requirements of users. From the perspective of grid stability and operational planning, this component is heavily influenced the generation and distribution of power supply. Consumption on the other hand, pertains to the whole quantity of electricity utilized within a certain timeframe, such as a day, week or month. While demand focuses on immediate energy needs, consumption offers a more comprehensive perspective on energy usage patterns over a period of time [8].

Peak demand is the maximum amount of electricity used at a specific period depending on diversity of consumption. Understanding peak demand is significant for power system planning and operation as it determines the capacity requirements of the grid to meet peak loads so that power outages and disruption can be avoided. Accurate peak demand forecasting aids in optimized the allocating resources in low-voltage distribution infrastructure [9].

Load profile is defined as the pattern of electrical energy consumption over time, reflecting how energy usage diverse throughout a day, week or year. Understanding demand characteristics is essential for utilities and energy providers to anticipate peak demand periods, optimize energy distribution, and enhance system reliability. Therefore, load profiles provide valuable insights for devising energy management strategies and plan infrastructure development.

Demand response is the ability of consumers to control their electricity consumption in response to changes of price or other incentives, aiming to reduce peak demand by redistribute load to off-peak periods for improving grid efficiency and costs savings. Various strategies like peak shaving and valley filling are employed in demand response programs to encourage consumers to lower energy usage during peak periods and shift it to times of lower demand.

2.1 Factor Influencing Electricity Demand

There are several factors influencing electricity demand in low-voltage distribution networks such as economic growth, population growth, climate change and technological advancements as illustrated in Fig. 1 [10, 11].

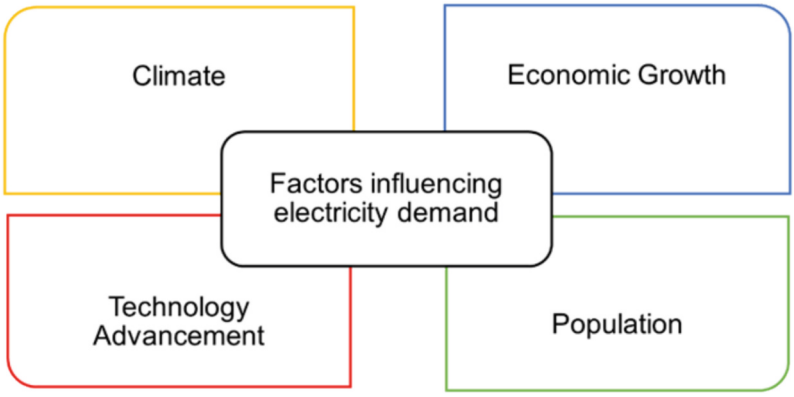


Fig. 1. Factor influencing electricity demand

As economies grow, the consumption of electricity increases significantly with rise in industrial activity, urbanization, and per capital income. Manufacturing sectors are the largest electricity consumers due to their energy-intensive nature. Meanwhile, urbanization amplifies demand through higher residential, commercial, and infrastructure energy needs. This demographic factor suggests that the more people there are, the more electricity required because urban areas and different demographic segments exhibiting varied energy needs [12]. Weather condition, since electricity is used in heating and cooling, seasonal changes and more broadly climate change are likely to influence this demand [13, 14]. Technological advancements have a dual impact, while digital devices increase demand, energy-efficient technologies help reduce it [15, 16]. Emerging technologies like electric vehicles and smart grids are shaping future demand, offering both threat and opportunities for energy management and integration [17].

2.2 Long-Term Trends in Electricity Demand

The long-term trends in electricity demand evolved by a complex interplay of demographic, economic, and technological factors. Generally, it increased over time as a result of population growth, industrialization, and urbanization. After World War II era, a significant surge in electricity demand due to industrial expansion and the invention of various sophisticated electronic devices have been witnessed. This evolvement can be considered as a game changer towards higher energy consumption as economies rebuilt and modernized. However, the growth in electricity demand has not been uniform. The 1970s and 1980s experienced a slowdown, primarily attributed to the oil price crisis that led to rationalization and conferred critical importance to the energy consumption patterns [18]. It is vital to analyze the connection between electricity demand and economic indicators such as GDP and industrial production for understanding these patterns. As economies grow, their energy needs are also increases and the electricity consumption often reflecting economic activity levels. However, this relationship does not remain simple. The transition from industrial economies to service based has led to a decrease in energy intensity, as services typically consume less energy compared to

manufacturing. Additionally, the availability of digital technologies has introduced new demand centers, altering traditional consumption patterns [19]. Moreover, technological advancements have also played a critical role in shaping electricity demand. Innovations in energy efficiency and the integration of renewable energy sources have moderated demand growth, even as economic activities expand. These milestones reflect that the demand of electricity is not stagnant as it rises with the growth of economy and also with the developments in technology.

2.3 Sectoral Trends in Electricity Consumption

Electricity consumption patterns vary significantly across sectors. The residential sector has seen a notable increase due to widespread use of appliances, air conditioning, and electric devices [20]. Conversely, the industrial sector has reduced its energy intensity through more efficient production processes and technologies, though overall demand remains high [21]. The commercial sector's electricity use has evolved with the growth of service industries and the digital economy, particularly driven by data centers and smart building technologies. Technological advancements and shifts in consumer behavior are influencing these sectoral patterns, with a move towards energy-efficient appliances and practices across all sectors.

2.4 Geographical Variations in Demand

Geographical variations in electricity demand are depend on various factors such as climate, population density, and economic growth [22]. Analysing these factors are important for planning electricity distribution more effective. For example, during winter and summer season, it affects heating and cooling needs resulting to seasonal variations in power consumption. Other than that, urbanized areas usually have more populations compared to rural areas, thus exhibit higher electricity consumption. Economically, a developed region shows higher consumption rates, reflecting improved living standards and greater use of electronic devices [23].

2.5 The Role of Peak Demand

Demand response of peak demand is essential for the grid stability and cost efficiency. Research study indicate that historical patterns of the peak demand now exhibit more complex patterns depending on time of day, day of the week, and season [24]. Furthermore, the integration of distributed renewable energy sources and power-intensive appliances adds complexity to peak demand forecasting and management [25]. Challenges include grid capacity constraints and balancing cost-effective asset management with demand response needs, requiring advanced analytical approaches for effective management.

2.6 Understanding Demand Elasticity

Demand elasticity of electricity refers to the responsiveness of electricity consumption to changes in price. Measuring elasticity of demand can be formulated as Eq. (1).

$$Elasticity\ of\ demand\ (E_{dl}) = \frac{Q_2 - Q_1}{P_2 - P_1} \times \frac{P_1}{Q_1} \tag{1}$$

From Eq. (1), when the absolute value is greater than 1, the elasticity of demand is said to be elastic and it indicates a relatively large change in quantity demanded in response to price changes. Meanwhile, when the absolute value is less than 1, it shows a small change in quantity demanded, and unit elasticity (when the absolute value is exactly 1, indicating proportionate changes in price and quantity demanded) [26]. Table 1 shows the relationship between price elasticity and changes in quantity demanded and also illustrated in Fig. 2.

Table 1. Relationship between price elasticity and changes in quantity demand

Types	Value of (E_{dl})		Example
Elastic	> 1	Big change in quantity demand Small change in prices	Food, household appliances, luxury car
Inelastic	< 1	Small change in quantity demand Big change in prices	Gasoline, electricity
Unit elastic	$= 1$	Proportionate change in quantity demand and prices	Electronic devices

Factors influencing demand elasticity include the availability of substitutes (more substitutes lead to higher elasticity) and the time horizon (longer time horizons tend to result in higher elasticity as consumers can adjust their behavior), and the necessity of the product (essential goods tend to have lower elasticity) [27].

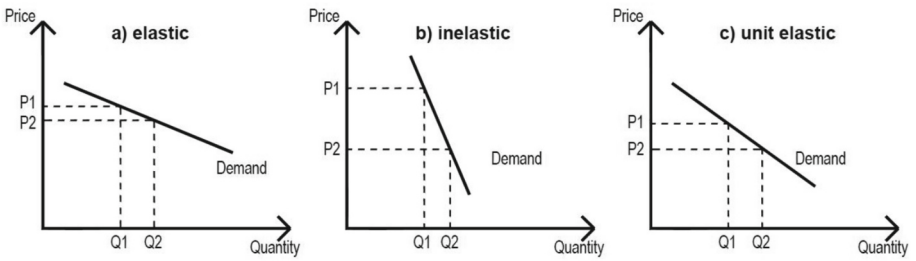


Fig. 2. Elasticity of demand (**a**) elastic, (**b**) inelastic and (**c**) unit elastic

2.7 Demand Elasticity and Low-Voltage Networks

Demand elasticity is a critical factor for managing low-voltage networks as it affects how consumers respond to changes in electricity prices, impacting load profiles and peak demand. Price-based demand response programs engage the demand elasticity in managing load variations and flattening peak demand. These programs can achieve a peak demand reduction of up to 15% and decrease the peak-to-valley span by nearly 47.33% [28]. This in turn enhances network reliability, reduces operational costs and optimizes the ability of the network to balance loads by reacting to price signals.

2.8 Implication for Network Investments

Understanding demand elasticity helps optimize network investments by influencing the timing and scale of upgrades. High demand elasticity indicates that consumers are sensitive to price changes hence reducing the need for immediate network expansions. Conversely, low demand elasticity result in condition where peak demand surpasses available capacity, necessitating more investments. By investigating demand elasticity trend, operators can prioritize investments, minimize overcapacity issues, and aspire network enhancements. This can result in cost savings by deferring or reducing infrastructure investments and expenditures.

2.9 Challenges and Opportunities

It is not easy to estimate the demand elasticity accurately for low-voltage networks due to the integration of advanced technologies like smart meters and electric vehicle charging station. Nonetheless, demand elasticity provides opportunities to enhance grid flexibility and resilience. It helps balance the variability of renewable energy sources and supports the integration of Distributed Renewable Energy Sources (DRES). The use of smart grid technologies and reinforcement learning algorithms can enhance microgrid functions and improve grid stability [29, 30]. Additionally, leveraging household flexibility and community-owned virtual power plants can further support renewable energy integration and trans-form distribution systems.

References

1. Ruifeng Z, Shuqing H, Donglin D, Xiaobing X, Anjiang L (2020) Low voltage power supply reliability evaluation of distribution network based on data quality governance. In: 2020 4th International conference on power and energy engineering (ICPEE). IEEE, pp 75–78
2. Bawa A, Uthman M, Shaibu FE, Oyewale KS (2019) Optimal sizing and sitting of distributed generation for power quality improvement of distribution network. Eur J Eng Res Sci 4:18–23
3. Sharma V, Garg P, Sharma A, Bharti N (2023) A comprehensive review: mitigating techniques for power variability due to integration of renewable energy sources with grid. J Electron Electr Eng
4. Wu Z, Bhat PK, Chen B (2023) Optimal configuration of extreme fast charging stations integrated with energy storage system and photovoltaic panels in distribution networks. Energies (Basel) 16:2385

5. Denis J, Florentin D (2024) Urban infrastructures' maturity and the age(s) of maintenance. In: Handbook of infrastructures and cities. Edward Elgar Publishing, pp 119–132
6. Gatta FM, Geri A, Maccioni M, Palazzoli A, Sancioni P (2024) Low voltage electric distribution network planning with demand control. Electr Power Syst Res 226:109950
7. Motahari S, Rahimpour MR (2024) Energy demand management. In: Encyclopedia of renewable energy, sustainability and the environment. Elsevier, pp 33–44
8. Atanane A, Benabbou L, El Ouafi A (2023) Electricity demand forecasting: a systematic literature review. In: 2023 14th International conference on intelligent systems: theories and applications (SITA). IEEE, pp 1–8
9. Donaldson DL, Browell J, Gilbert C (2024) Predicting the magnitude and timing of peak electricity demand: a competition case study. IET Smart Grid 7:473–484
10. Su D, Tan B, Zhang A, Hou Y (2023) Analysis of the influencing factors of power demand in Beijing based on the LMDI model. Sustainability 15:7913
11. Liu Q et al (2023) Study on the spatial and temporal evolution system of the factors influencing electricity consumption in China. In: Zhang K, Grigoras G (eds) International conference on mechatronics and intelligent control (ICMIC 2023). SPIE, p 82
12. Rodríguez-Caballero CV (2022) Energy consumption and GDP: a panel data analysis with multi-level cross-sectional dependence. Econom Stat 23:128–146
13. Romitti Y, Sue Wing I (2022) Heterogeneous climate change impacts on electricity demand in world cities circa mid-century. Sci Rep 12:4280
14. Vu DH, Muttaqi KM, Agalgaonkar AP (2014) Assessing the influence of climatic variables on electricity demand. In: 2014 IEEE PES general meeting | conference & exposition. IEEE, pp 1–5
15. Forsythe CR, Gillingham KT, Michalek JJ, Whitefoot KS (2023) Technology advancement is driving electric vehicle adoption. In: Proceedings of the national academy of sciences, vol 120
16. Chen W (2022) Guest editorial for the special issue on: energy technology advancements. Int J Energy Res 46:18928–18928
17. Naim A, Vashist D (2022) Technological advancements for reduced charging time of electric vehicle batteries: a review. Presented at the
18. Osisanwo B, Babasanya A, Ogede J, Siyanbola A (2023) Effects of electricity demand on economic growth in Nigeria. NIU J Soc Sci 9:7–14
19. Pesantez JE, Li B, Lee C, Zhao Z, Butala M, Stillwell AS (2023) A comparison study of predictive models for electricity demand in a diverse urban environment. Energy 283:129142
20. Archana (2024) Investigating electricity usage profile for residential consumers in India. Strategic Planning for Energy and the Environment, pp 499–516
21. Xu N (2023) Energy consumption forecasting: a trend analysis based on global data from 2000–2020. In: International conference on mathematics and machine learning. ACM, New York, NY, USA, pp 146–151
22. Zhang Y, Li J (2020) Geographic variation in Medicare per capita spending narrowed from 2007 to 2017. Health Aff 39:1875–1882
23. Nikishin A, Parshilkina A, Tristanov A (2023) Regional power system consumption analysis with data mining techniques. Presented at the
24. Van Orden E, Simović U (2023) Managing peak load in the past, present, and future: challenges and opportunities for a flexible energy future. Clim Energy 39:15–20
25. El Gohary F, Stikvoort B, Bartusch C (2023) Evaluating demand charges as instruments for managing peak-demand. Renew Sust Energ Rev 188:113876
26. Cvitković D, Lolić Čipčić M (2023) Analysis of the elasticity of demand for bread. Elektronički zbornik radova Veleučilišta u Šibeniku 17:67–78
27. Kawagoe BR (2022) Analysis of price-demand elasticity of electric energy in smart electrical grids. J Eng Res 2:2–12

28. Kansal G, Tiwari R (2023) Elasticity modelling of price-based demand response programs considering customer's different behavioural patterns. Sustain Energy Grids Netw 36:101244
29. Kumar PP et al (2024) Grid stability enhancement through machine learning-driven control strategies in renewable energy integration. In: 2024 12th International conference on smart grid (icSmartGrid). IEEE, pp 317–321
30. Onu P, Pradhan A, Madonsela NS (2024) Autonomous microgrids optimization using reinforcement learning: applications, challenges and prospects. In: 2024 1st International conference on smart energy systems and artificial intelligence (SESAI). IEEE, pp 1–6

Design and Simulation of an Interleaved Boost Converter to Stabilize the Output Voltage under Varying Condition

Andre Dimas Yogatama, Febiola Sabrina, and Arya Kusumawardana(✉)

State University of Malang, Malang, East Java 65145, Indonesia
arya.kusumawardana.ft@um.ac.id

Abstract. Recently, Interleaved Boost Converters (IBCs) were one of the most helpful breakthroughs in DC-to-DC voltage boost applications for large power. Despite the many benefits of this topology, the multi-phase concept has its own difficulties in terms of controller design and implementation. The methods used for controller design in brief are literature review, simulation tool selection, model design, testing, analysis, and report. This research presents an alternative controller design to ease the implementation. A single op-amp PI controller is proposed to maintain the voltage stability of the IBCs output. There are three tests to determine the performance of the designed controller. The first test is by providing a constant input voltage and load. The second test is done by changing the load value with a constant input voltage. The next is by changing both the input voltage and the load. Based on the test results, with variations in load and input voltage, the performance of the converter with the proposed controller remains stable, characterized by a steady state error close to 0 in a limited time.

Keywords: Interleaved Boost Converter · Voltage Control · PI Controller · Op-amp

1 Introduction

Interleaved boost converters (IBCs) are critical in applications requiring the elevation of low voltage inputs to high voltage outputs, particularly in renewable energy systems [1, 2]. They offer several advantages over conventional boost converters, including improved efficiency, reduced size, and electromagnetic emissions, as well as enhanced transient response and reliability [3, 4]. These converters are especially beneficial in high-power applications such as multimegawatt wind energy conversion systems [5] and high-power step-up and power factor correction applications [6].

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 402–417, 2026.
https://doi.org/10.1007/978-981-95-9233-3_29

IBCs significantly enhance renewable energy systems like photovoltaic (PV), wind, and hydro power by addressing variability and low voltage issues. IBCs increase voltage for grid integration or direct use while reducing input current ripple, improving efficiency and reliability [7, 8]. Although IBCs are often linked to PV systems for efficiently managing low voltage outputs and maximizing power via Maximum Power Point Tracking (MPPT), they also benefit wind and hydro power by stabilizing variable outputs [9, 10]. The multi-phase nature of IBCs reduces electromagnetic interference (EMI) and downsizes power electronics, which is advantageous for all renewable applications [11, 12]. Overall, IBCs enhance renewable energy systems by providing high efficiency, smaller component size, and stable voltage output, validating their adaptability and efficiency across various renewable applications [7, 8, 11].

Interleaved boost converters require precise controller design to manage nonlinearity and ensure stability under varying conditions [13]. These converters, used in renewable energy applications, demand accurate control for desired output voltage and minimized power losses and ripple [13, 14]. Control strategies like fuzzy logic controllers (FLC), proportional-integral (PI) controllers, and model predictive control (MPC) offer distinct advantages in performance and efficiency [13–15]. FLCs minimize error and eliminate output ripples, while PI controllers enhance efficiency and reduce ripple current [14]. MPC effectively distributes power among boost channels and integrates with existing algorithms for specific applications. Adaptive output feedback controllers address system non-linearity, load uncertainty, and configurations lacking current sensors [16].

Advancements in interleaved boost converter voltage control have emphasized enhancing efficiency, response time, and reliability. Sakasegawa et al. [17] introduced an LQI control method that improves response to voltage step changes and maintains stable output under load variations, outperforming PI control. Wu et al. in 2021 [18] proposed a fuzzy control algorithm for parallel current sharing in four-phase interleaved boost converters, enhancing response speed, reducing overshoot, output current ripple, and voltage. An FCS-MPC was developed for effective power distribution and switching frequency stabilization [15]. Keskin et al. in 2021 introduced a structured H-infinity synthesis framework for designing robust Type-III controllers, ensuring stability against model uncertainties and improving voltage regulation under disturbances [19].

In tandem with advancements in controller development, hardware design innovations have tackled challenges including component stress, thermal management, and electromagnetic interference (EMI) [20]. The integration of advanced semiconductor devices, particularly wide-bandgap materials such as Silicon Carbide (SiC) and Gallium Nitride (GaN), has facilitated higher switching frequencies and enhanced overall converter performance [21, 22]. Additionally, innovative magnetic component designs and PCB layout techniques have played a crucial role in achieving more compact and efficient converter implementations [23, 24].

Designing effective control systems for interleaved boost converters is challenging due to their complex multi-phase dynamics, often leading to suboptimal performance across the entire operating range [25–27]. The need for multiple sensors and sophisticated algorithms can increase costs and reduce the reliability of interleaved boost converters [28, 29]. Despite these advancements, there remains a persistent need for simple, cost-effective control solutions that can deliver robust performance without sacrificing

reliability or ease of implementation. This research presents a novel approach to voltage control of interleaved boost converters using a single operational amplifier (op-amp) as a PI controller. The proposed solution offers several advantages over existing methods:

- Simplicity: By utilizing a single op-amp, the control circuitry is significantly simplified, reducing component count and potential points of failure.
- Cost-effectiveness: The minimalist approach translates to lower implementation costs, making it an attractive option for budget-constrained applications.
- Analog implementation: Unlike digital control solutions, the proposed analog controller offers instantaneous response and does not suffer from sampling and quantization effects.
- Ease of tuning: The PI parameters can be easily adjusted using standard passive components, facilitating rapid prototyping and optimization.
- Compatibility: The single op-amp PID controller can be readily integrated into existing interleaved boost converter designs with minimal modifications.

The research presented in this paper builds upon the foundational work in power electronics and control theory while offering a novel perspective on simplifying control structures for complex power conversion systems. By demonstrating the efficacy of a single op-amp PI controller for interleaved boost converters, this study opens new avenues for cost-effective and reliable control solutions in renewable energy applications.

2 Research Method

The research on the Interleaved Boost Converter system refers to several relevant previous studies. The findings from these studies have been summarized in a table to analyze the obtained results and the differences in control systems used. The analysis of these studies can be seen in Table 1.

This simulation is designed to create an experiment that closely mirrors real-world conditions, focusing on the development of an Interleaved Boost Converter with the function of increasing the voltage from a generator. To achieve this goal, a methodical and systematic approach is necessary to ensure that the research is well-planned and executed, as in Fig. 1.

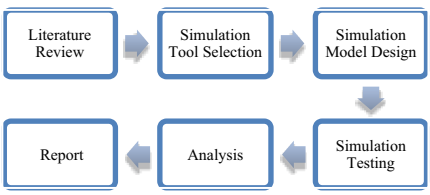


Fig. 1. Research flowchart

Table 1. Table of previous studies

Research title	Research results	Control system
Design and development of an Arduino-based interleaved boost converter [30]	In this study, the interleaved boost converter research method using Arduino showed a significant reduction in input current ripple and output voltage of the boost converter. Simulation and testing of the interleaved boost converter with various duty cycle values demonstrated that the reduction in the input current ripple and output voltage is considerably smaller	In this study, the controller used was an Arduino-based controller
Design and implementation of interleaved boost converter [31]	This study investigates the design and simulation of an interleaved boost converter for sustainable non-conventional energy sources. The two-phase interleaved DC/DC boost converter has several advantages over a conventional boost converter, including higher efficiency, faster dynamics, and minimal current ripple. The developed converter is capable of increasing the output voltage up to 400 V	This study uses the state space averaging method to analyze the interleaved boost converter. Using this method, mathematical model equations are required for the two-phase interleaved boost converter. The operation of the converter is explained with the help of four switching modes
Design and analysis of a closed-loop boost converter [32]	In this study, the results of the closed-loop boost converter circuit with a constant load of 129.96 Ω and different duty cycle values showed that the average output voltage produced was 36.082 Volts, and it remained stable at the desired point. By using a fuzzy logic controller, the transient response was improved by reducing the average settling time (ts) to 30.50281 ms, the rise time (tr) to 1.3761 ms, and the peak time (tp) to 4 ms	This study uses a fuzzy logic controller in the closed-loop boost converter circuit to regulate the MOSFET duty cycle

2.1 Model Design

To make a device, the design of the device will be needed. The overall model of the interleaved boost converter design can be seen in Fig. 2. The design model utilizes Simscape Electrical, a toolbox that provides simulations closely aligned with real-world conditions.

This model encompasses all system components, including the control system and block diagrams. By using Simscape Electrical, the design ensures accurate representation and analysis of the entire system, facilitating effective modeling and optimization of the Interleaved Boost Converter.

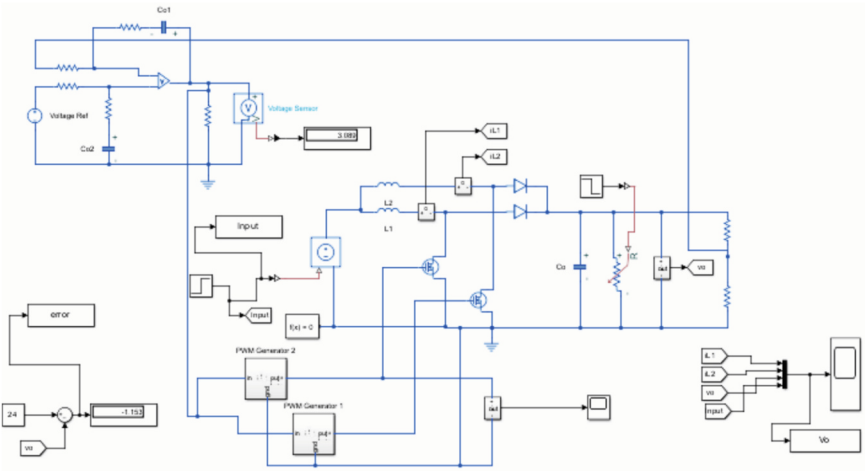


Fig. 2. Design model of the interleaved boost converter system

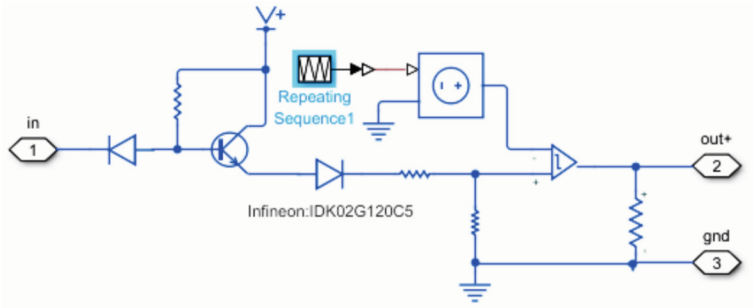


Fig. 3. Voltage-to-duty cycle converter (IC LM5032)

The image above illustrates the converter design used in this research, which is based on the LM5032 IC. This IC is designed to increase voltage and manage the relationship between input voltage and duty cycle. The converter functions to determine the output signal and frequency through a comparator. The LM5032 IC plays a critical role in regulating the boost converter's performance by adjusting the duty cycle to achieve the desired output voltage, as in Fig. 3.

Figures 4 and 5 illustrate the design of the PI controller (Proportional-Integral), which incorporates both Kp (proportional gain) and Ki (integral gain). To determine the

optimal PI controller settings, a process of fine-tuning will be employed. This involves adjusting the input values and iterating through different settings until the smallest RMS value and error are achieved. The input values for both Kp and Ki will range from 0.1 to 1, and the goal is to minimize output deviations and achieve stable control performance.

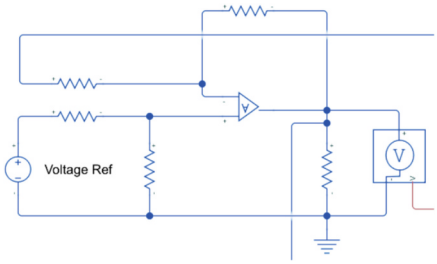


Fig. 4. Design of a proportional (P) controller

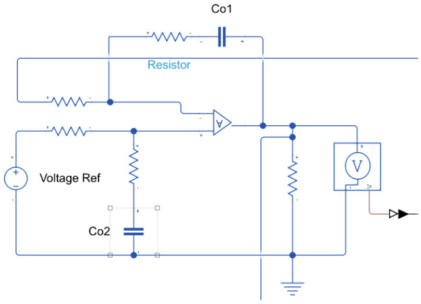


Fig. 5. Design of an integral controller

2.2 Converter Design

One important factor when designing IBCs is determining the value of passive components used. Some parameters that need to be determined include Duty Cycle as in Eq. (1), output current, as in Eq. (2), input current, as in Eq. (3), ripple voltage, as in Eq. (4), ripple current, as in Eq. (5), inductor value, as in Eq. (6), and capacitor, as in (7). These parameters can be determined as follows [33]:

Duty Cycle:

$$\mathrm{D} = \frac{1}{1-\mathrm{D}}\mathrm{D} = \frac{1}{1-\mathrm{D}} \tag{1}$$

Output Current:

$$I = \frac{P_{out}}{V_{out}} \tag{2}$$

Input Current [34]:

$$I_{in} = \frac{Pin}{Vin} \tag{3}$$

Ripple Voltage:

$$\Delta V = 0.01 \times V_{out} \tag{4}$$

Ripple Current:

$$\Delta I = 0.05 \times I_{in} \tag{5}$$

The minimum inductor value $L_{min}L_{min}$ [35]:

$$L_{min} = \frac{D_{min}\ (1 - D_{min})^2\ R_{max}}{2f_{min}} \tag{6}$$

Capacitor Output:

$$C = \frac{\mathrm{D}\ \times V_{out}}{F_{sw} \times\ \Delta \mathrm{V}\ \times\ R_{out}} \tag{7}$$

2.3 Result and Discussion

IBCs are nonlinear systems that require a linearization mechanism to design the controller, especially PID. However, in this study, the PI controller is designed with a fine-tuning mechanism using simulation. Simulations performed using MATLAB software were conducted with the parameters presented in Table 2 (based on Eq. (1)–(7)).

Table 2. Interleaved boost converter parameters

Key parameters of the interleaved boost converter			
Parameter	Symbol	Value	Unit
Input voltage	V_{in}	12	V
Output voltage	V_{out}	24	V
Power	P	200	W
Capacitor	C_1	100	uF
Capacitor	C_2	100	uF
Inductor	L_1	50	uH
Inductor	L_2	50	uH
Output load resistor	R	2,89	Ω
Frequency *switching* Mosfet	f	100	kHz

The device testing parameters were carried out through several tests: first, by varying the input voltage; second, by varying the load; and third, by conducting a comprehensive overall test, as presented in Table 3

Determining the values of Kp and Ki is also necessary to ensure that the control system performs adequately, as presented in Table 3. The values of Kp and Ki were obtained using the fine-tuning method.

Table 3. PI control parameters

Parameter	Value
Kp	0.1
Ki	0.3
Frequency switching	100 kHz

2.4 Input Voltage Variation Testing

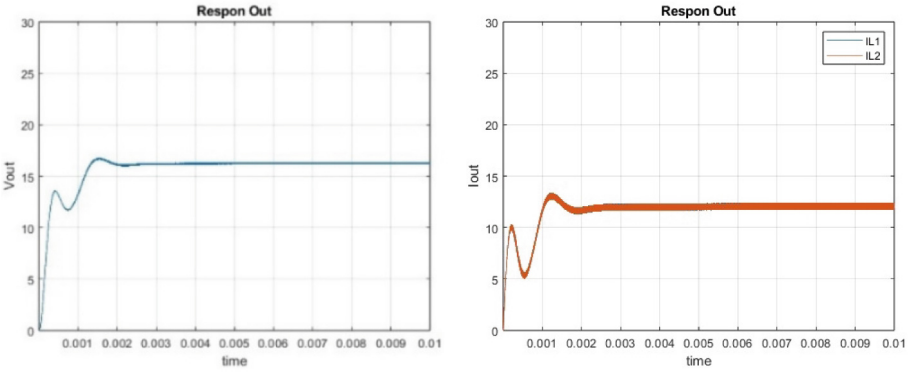
Voltage variation testing was conducted by applying input voltages ranging from 5 V to 12 V with a fixed load value of 2.89 Ω. The testing duration ranged from 0.001 to 0.01 s. Prior to this test, the values of Kp and Ki were determined to optimize the control system for the simulation experiment. The results are presented in the form of response graphs that illustrate the effects of varying the input voltage. The test data for voltage variation can be found in Table 4.

Table 4. Data from voltage variation testing

Testing with input voltage variation						
Input voltage (V)	Load (Ohm)	Stady state error	Overshoot	Rise time (s)	Settling time (s)	RMSE
5	2.89	7.73	–	–	–	6.62
6	2.89	4.38	–	–	–	5.78
7	2.89	1.04	–	–	–	3.83
8	2.89	0.01	–	0.004	0.004	3.35
9	2.89	0.02	0.05	0.004	0.004	3.05
10	2.89	0.02	1.85	0.004	0.004	2.80
11	2.89	0.02	4.16	0.004	0.004	2.65

Simulation results from MATLAB with input voltages ranging from 5 V to 11 V showed that the error value was relatively high at an input voltage of 5 V and decreased as the input voltage increased. The RMSE value also decreased with increasing input voltage, starting from a high RMS value of 6.62 at 5 V. At an input voltage of 11 V, the RMS value was significantly lower at 2.62. For tests with input voltages of 5 V, 6 V, and 7 V, the Vout response did not reach the desired setpoint of 24 V. The response of the output voltage and both inductor currents is shown in Fig. 6.

The current ripple produced by the system is 5%, indicating effective interleaving operation in reducing current ripple and achieving balanced load distribution. The oscillation frequency and amplitude are stable, ranging from approximately 11.8 to 12.4, as shown in Fig. 7.



(a) (b)

Fig. 6. System responses at 5 V input voltage; (**a**) output voltage response; (**b**) inductor current response

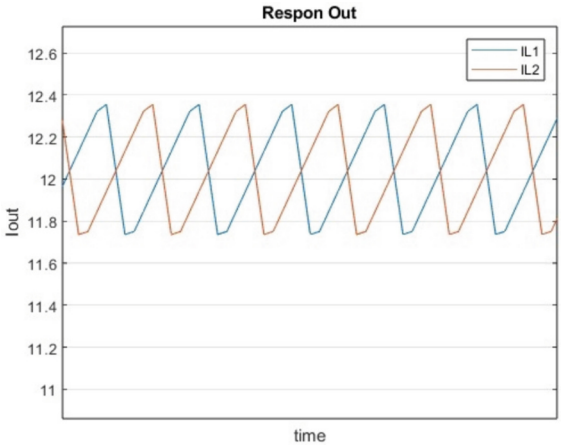


Fig. 7. Current ripple in the 5 V testing

2.5 Load Variation Testing

This test involves evaluating the system with a constant input voltage of 12 V and varying loads, ranging from 24 Ω to 2.89 Ω. The test uses the same Kp and Ki control system values as in the previous tests. The results of the load variation testing can be found in Table 5. MATLAB testing results with an input voltage of 12 V and varying loads from 24 Ω to 2.89 Ω reveal that as the resistance or load decreases, both RMSE and overshoot values decrease. This indicates that with higher loads, the system becomes more stable (with lower RMSE values). Lower RMSE values suggest improved system performance, while reduced overshoot indicates greater accuracy.

Table 5. Results of load variation testing

Load variation testing						
Input (V)	Load (Ohm)	Stady state error	Overshoot	Rise time (s)	Settling time (s)	RMSE
12	24	0.003	13.3	0.005	0.005	4.22
12	12	0.02	12.24	0.005	0.005	3.28
12	8	0.02	11.2	0.004	0.004	2.96
12	6	0.02	10.25	0.003	0.003	2.77
12	4.8	0.02	9.3	0.003	0.003	2.68
12	4	0.02	8.5	0.003	0.003	2.64
12	3.4	0.02	7.48	0.004	0.004	2.65
12	2.89	0.02	6.45	0.005	0.005	2.60

Simulation results from MATLAB, as shown in Fig. 8, demonstrate that applying a reference voltage of 12 V yields a constant output voltage (Vout) of 24 V. The current ripple shows oscillations ranging from 9.2 to 10.2, indicating that the current remains stable with a ripple of 5% or 1.2 V Fig. 9b.

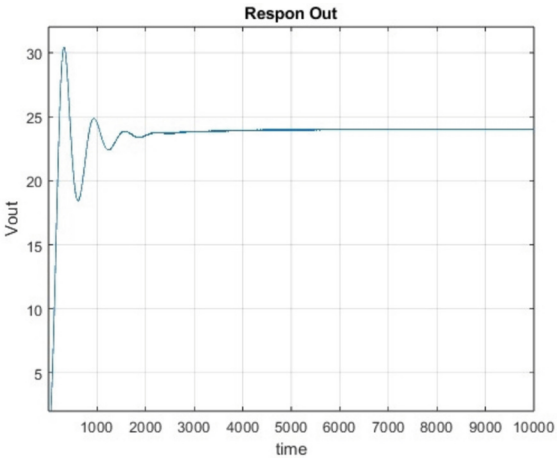
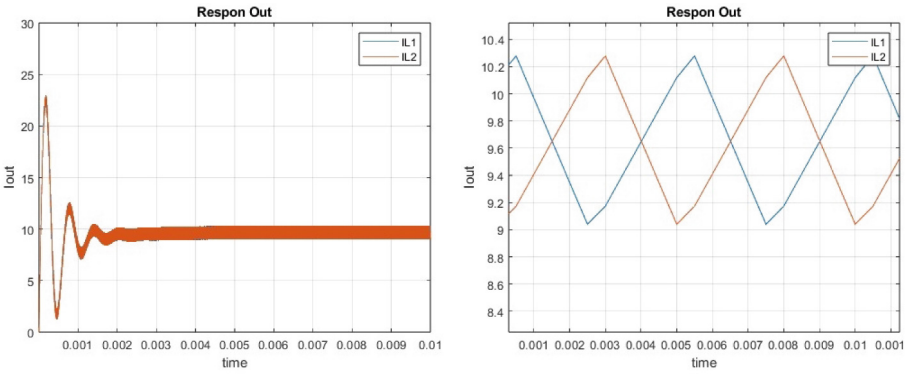


Fig. 8. Vout response with a load variation of 2.89 Ω



(a) (b)

Fig. 9. Inductor current responses; (**a**) output inductor current; (**b**) inductor current ripple

2.6 Load and Input Voltage Variation Testing

The final testing involves applying different input values at different times with varying load variables. The simulation runs for 0.02 s and includes two test trials. This final test aims to assess the system's behavior under various conditions during the simulation period. Testing with an input voltage of 9-11 V and loads ranging from 24 Ω to 2.89 Ω involves initially applying 9 V for a period, then changing to 11 V after 0.007 s with a load of 24 Ω. At 0.014 s, the load is adjusted to 2.89 Ω. The load variation is tested to observe the system's response and determine if it can maintain a consistent output voltage.

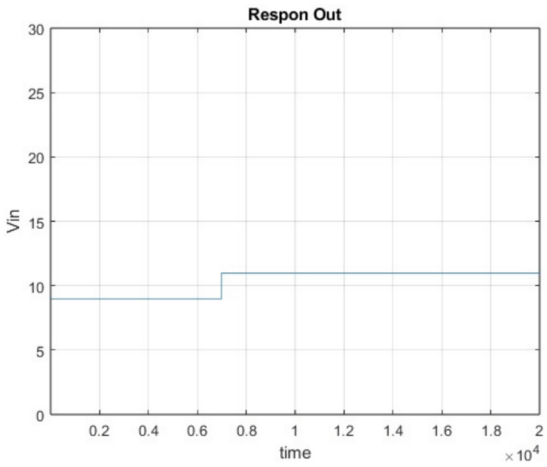


Fig. 10. Results of input response from 9 V to 11 V

Figure 10 shows the overall testing by applying different input voltages. From 0 to 0.007 s, an input voltage of 9 V is applied, and after this period, the voltage is changed to 11 V. The test is conducted over a total duration of 0.02 s.

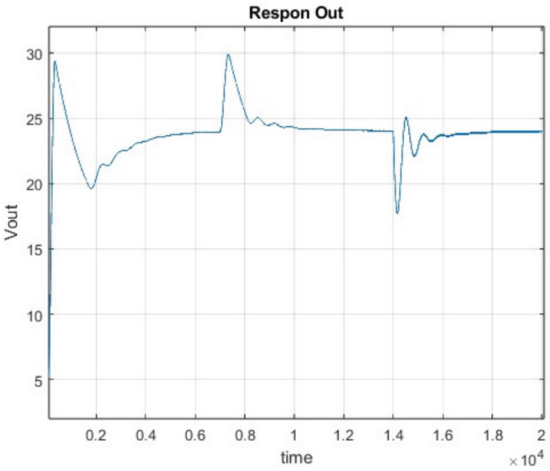


Fig. 11. Vout response from input variation testing from 9 V to 11 V

The Vout response graph as shown in Fig. 11 indicates that the system experiences overshoot and oscillation before reaching a steady-state condition at 24 V, suggesting potential instability and the need for further tuning of system parameters to achieve a faster and more stable response. The graph shows that the system takes a relatively long time to stabilize and is affected by external disturbances causing fluctuations in the output voltage. At 0.007 s, voltage changes lead to overshoot oscillations due to high current conditions. Additionally, at 0.014 s, a load change from 24 Ω to 2.89 Ω results in oscillations, marked by an increase in current.

The output current graph illustrates the system's response to changes in input voltage and load, as in Fig. 12. When the input voltage is increased, the output current rises in accordance with basic electrical principles. However, during drastic load changes, the system experiences instability, characterized by current spikes and oscillations. This suggests factors such as inductance, capacitance, or excessively high system gain that affect system stability. Despite this, the system eventually reaches a stable condition at a higher current value.

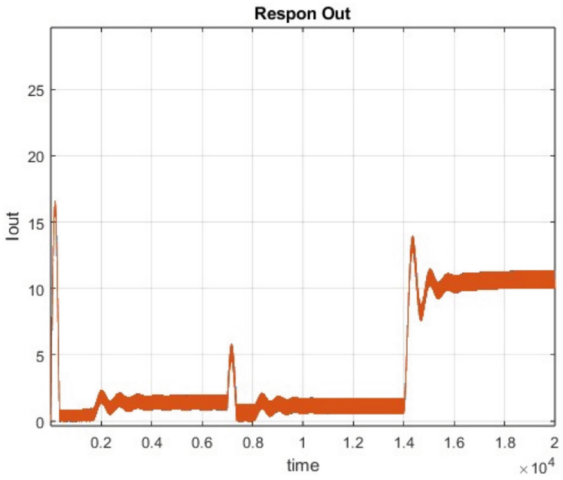


Fig. 12. Response to changes in output current in the system

3 Conclusion

Based on the test results, it can be seen that the PI controller concept by utilizing one Op-amp can be effectively used as a voltage stabilizer. In the first test, there are interesting things when the input voltage is adjusted from 5 V - 8 V. There are several parameters such as overshoot, rise time and settling time that are undefined. This is because the system output is not able to reach its reference voltage. It can be said that the minimum value of input voltage that can be used is at 8 V. When the input voltage starts at 9 V, the system can reach the reference voltage characterized by a smaller steady state error value. Testing load changes do not have a significant impact on system performance, except for the overshoot value. The smaller the resistance value used, the smaller the overshoot value. Voltage and current spikes occur when there is a sudden change in input voltage and load. However, for a limited time the output voltage can approach the reference value. The results of each test show that the system with minimum hardware design is robust because it can maintain the system response when a disturbance occurs. However, current spikes that occur due to sudden load changes still need to be resolved. Therefore, it is necessary to design a current limiter that does not significantly complicate the controller to limit the inductor current value so as not to exceed the capabilities of the components used.

References

1. Mittle A, Singh RK, Sarath Chandra J (2019) A new interleaved high step-up DC-DC converter. In: 2019 IEEE students conference on engineering and systems (SCES). IEEE, pp 1–5. https://doi.org/10.1109/SCES46477.2019.8977225
2. Valarmathy AS, Prabhakar M (2024) High gain interleaved boost-derived DC-DC converters—a review on structural variations, gain extension mechanisms and applications. E-Prime—Adv Electr Eng, Electron Energy 8:100618. https://doi.org/10.1016/j.prime.2024.100618

3. Khosroshahi A, Abapour M, Sabahi M (2015) Reliability evaluation of conventional and interleaved DC–DC boost converters. IEEE Trans Power Electron 30:5821–5828. https://doi.org/10.1109/TPEL.2014.2380829
4. Scofield J, McNeal S, Jordon B, Kosai H, Ray B (2011) Studies of interleaved DC-DC boost converters with coupled inductors. https://doi.org/10.21236/ADA542736
5. Aguirre M, Kouro S, Rodriguez J, Abu-Rub H (2015) Model predictive control of interleaved boost converters for synchronous generator wind energy conversion systems. In: 2015 IEEE international conference on industrial technology (ICIT). IEEE, pp 2295–2301. https://doi.org/10.1109/ICIT.2015.7125436
6. Makee AH, Al-Omari AH (2023) Employing various topologies to improve the interleaved boost converters performance. Indones J Electr Eng Comput Sci 31:54–67. https://doi.org/10.11591/ijeecs.v31.i1.pp54-67
7. Newlin DJS, Ramalakshmi R, Rajasekaran S (2013) A performance comparison of interleaved boost converter and conventional boost converter for renewable energy application. In: 2013 International conference on green high performance computing (ICGHPC). IEEE, pp 1–6. https://doi.org/10.1109/ICGHPC.2013.6533924
8. Gupta R, Gauta S, Kumar M (2022) Comparative analysis of boost and interleaved boost converter for PV application. In: 2022 2nd International conference on emerging Frontiers in electrical and electronic technologies (ICEFEET). IEEE, pp 1–6. https://doi.org/10.1109/ICEFEET51821.2022.9848264
9. Kumar M, Panda G, Shiva Krishna Rao DVK (2022) Analysis of conventional and interleaved boost converter with solar photovoltaic system. In: 2022 International conference on intelligent controller and computing for smart power (ICICCSP). IEEE, pp 1–6. https://doi.org/10.1109/ICICCSP53532.2022.9862351
10. Tang C-Y, Chen Y-T, Chen Y-F, Chen Y-M, Chang Y-R (2013) Multi-mode interleaved boost converter for photovoltaic power systems with low-voltage ride-through capability. In: 2013 IEEE energy conversion congress and exposition. IEEE, pp 1096–1101. https://doi.org/10.1109/ECCE.2013.6646827
11. Rahavi JSA, Kanagapriya T, Seyezhai R (2012) Design and analysis of interleaved boost converter for renewable energy source. In: 2012 International conference on computing, electronics and electrical technologies (ICCEET). IEEE, pp 447–451. https://doi.org/10.1109/ICCEET.2012.6203850
12. Ramesh A, Kumar MS, Sekhar OC (2016) Interleaved boost converter fed with PV for induction motor/agricultural applications. Int J Power Electron Drive Syst (IJPEDS) 7:835. https://doi.org/10.11591/ijpeds.v7.i3.pp835-853
13. Rexy I, Mary N, Starlin J, Elizabeth R (2022) Interleaved boost converter with fuzzy logic controller. In: 2022 First international conference on electrical, electronics, information and communication technologies (ICEEICT). IEEE, pp 1–7. https://doi.org/10.1109/ICEEICT53079.2022.9768492
14. Devi MA, Valarmathi K, Mahendran R (2014) Ripple current reduction in interleaved boost converter by using advanced PWM techniques. In: 2014 IEEE international conference on advanced communications, control and computing technologies. IEEE, pp 115–119. https://doi.org/10.1109/ICACCCT.2014.7019291
15. Aguirre M, Kouro S, Rodriguez J, Abu-Rub H (2015) Model predictive control of interleaved boost converters for synchronous generator wind energy conversion systems. In: 2015 IEEE international conference on industrial technology (ICIT). IEEE, pp 2295–2301. https://doi.org/10.1109/ICIT.2015.7125436
16. Salhi B, El Fadil H, Ahmed Ali T, Magarotto E, Giri F (2015) Adaptive output feedback control of interleaved parallel boost converters associated with fuel cell. Electr Power Compon Syst 43:1141–1158. https://doi.org/10.1080/15325008.2014.995280

17. Sakasegawa E, Watanabe S, Shiraishi T, Haga H, Kennel RM (2024) A novel LQI control technique for interleaved-boost converters. World Electr Veh J 15:343. https://doi.org/10.3390/wevj15080343
18. Wu W, Chen J, Chen Z, Li Z, Xiao C (2021) Parallel current sharing control algorithm of four phase interleaved boost converter. In: 2021 China automation congress (CAC). IEEE, pp 1860–1864. https://doi.org/10.1109/CAC53003.2021.9728115
19. Keskin R, Aliskan I, Daş E (2021) Robust structured controller synthesis for interleaved boost converters using an H_∞ control method. Trans Inst Meas Control 43:3169–3180. https://doi.org/10.1177/01423312211019560
20. Blaabjerg F, Yang Y, Ma K, Wang X (2015) Power electronics—the key technology for renewable energy system integration. In: 2015 International conference on renewable energy research and applications (ICRERA). IEEE, pp 1618–1626. https://doi.org/10.1109/ICRERA.2015.7418680
21. Millan J, Godignon P, Perpina X, Perez-Tomas A, Rebollo J (2014) A survey of wide bandgap power semiconductor devices. IEEE Trans Power Electron 29:2155–2163. https://doi.org/10.1109/TPEL.2013.2268900
22. Jones EA, Wang FF, Costinett D (2016) Review of commercial GaN power devices and GaN-based converter design challenges. IEEE J Emerg Sel Top Power Electron 4:707–719. https://doi.org/10.1109/JESTPE.2016.2582685
23. McLyman CWT (2004) Transformer and inductor design handbook. CRC Press. https://doi.org/10.1201/9780203913598
24. Biglarbegian M, Shah N, Mazhari I, Enslin J, Parkhideh B (2016) Design and evaluation of high current PCB embedded inductor for high frequency inverters. In: 2016 IEEE applied power electronics conference and exposition (APEC). IEEE, pp 2998–3003. https://doi.org/10.1109/APEC.2016.7468290
25. Zurbriggen IG, Anun M, Ordonez M (2014) Dynamic physical limits of boost converters: a benchmarking tool for transient performance. In: 2014 IEEE energy conversion congress and exposition (ECCE). IEEE, pp 4118–4124. https://doi.org/10.1109/ECCE.2014.6953962
26. Nikhar AR, Apte SM, Somalwar R (2016) Review of various control techniques for DC-DC interleaved boost converters. In: 2016 International conference on global trends in signal processing, information computing and communication (ICGTSPICC). IEEE, pp 432–437. https://doi.org/10.1109/ICGTSPICC.2016.7955340
27. Zhang S, Yu X (2012) Control strategy to achieve minimum/zero input current ripple for the interleaved boost converter in photovoltaic/fuel cell power conditioning system. In: 2012 IEEE energy conversion congress and exposition (ECCE). IEEE, pp 4301–4306. https://doi.org/10.1109/ECCE.2012.6342237
28. Liu G, Wang M, Zhou W, Wu Q, Fu Y (2021) A Sensorless current balance control method for interleaved boost converters based on output voltage ripple. IEEE Trans Power Electron 36:7138–7149. https://doi.org/10.1109/TPEL.2020.3037650
29. Mazumder S, Tolani S, Joshi S (2022) Low-cost single-sensor current sharing scheme for two-phase DC-DC boost converter. In: 2022 IEEE 10th power India international conference (PIICON). IEEE, pp 1–6. https://doi.org/10.1109/PIICON56320.2022.10044999
30. Mazta MA, Samosir AS, Haris A (2016) Design and development of an Arduino-based interleaved boost converter. ELECTRICIAN–Jurnal Rekayasa dan Teknologi Elektro 10. https://doi.org/10.23960/elc.v10n1.190
31. Shenoy KL, Nayak CG, Mandi RP (2017) Design and implementation of interleaved boost converter. https://doi.org/10.21817/ijet/2017/v9i3/170903S076
32. Amalia RN, Kusuma W, Yuniq HA (2023) Design and analysis of a closed-loop boost converter, vol 6, pp 130–139

33. Pirashanthiyah L, Edirisinghe HN, De Silva WMP, Bolonne SRA, Logeeshan V, Wanigasekara C (2024) Design and analysis of a three-phase interleaved DC-DC boost converter with an energy storage system for a PV system. Energies 17:1–14. https://doi.org/10.3390/en17010250
34. Assyidiq MA, Winardi B, Andromeda T: Design of a boost converter with voltage feedback for solar panels (2017)

Drone Decentralized Scheduling for Rural Area Monitoring

Muhammad Zakiyullah Romdlony[1], Rashad Abul Khayr[1](✉), Porman Pangaribuan[1], Yul Yunazwin Nazaruddin[2], and Tua Agustinus Tamba[3]

[1] School of Electrical Engineering, Telkom University, Bandung, Indonesia
rashadaka@student.telkomuniversity.ac.id
[2] Faculty of Industrial Technology, Institut Teknologi Bandung, Bandung, Indonesia
[3] Department of Electrical Engineering, Parahyangan Catholic University, Bandung, Indonesia

Abstract. Unmanned Aerial Vehicles (UAV) have become more popular due to their wide array of applications, such as monitoring tasks. However, the drone's power and the environment constrain this task. Reducing the drone's time and distance can mitigate the power issue. A decentralized scheduling method can reduce the impact on the environment. The chaos-based bidding consensus method demonstrates a balanced approach to task scheduling, reducing the total time required and the total distance traveled. Compared to the greedy method, the chaos-based bidding method requires 16.95% less time and 9.69% less distance for the highest requirement. This approach also comes with fewer overloading tasks for a single agent, making this method a good choice for a resource-critical environment. Coupled with the decentralized nature of the chaos-based bidding method, the proposed method can be used for drone task scheduling, especially in a limited network and resource-critical environments like a disaster-affected area or an underdeveloped rural area.

Keywords: Decentralized Scheduling · UAV · Chaos-Based Bidding Consensus

1 Introduction

Unmanned Aerial Vehicles (UAV) have become more popular due to its wide array of application [1]. UAV more popularly called drone and its ability to do many types of tasks comes from its mobility and flexibility [2, 3]. One of the types that the drone can do is monitoring tasks. The drone was used for field observation, patrol, or photographing unhampered by the terrain due to its ability to fly, but not without disadvantages of its own. Drones require a complex control system and a lightweight so they can fly. This limits the amount of battery despite the power needed to fly the drone [4, 5]. To combat this problem, drone scheduling can be used to limit drone usage [6, 7]. For the monitoring problem, various research was done to minimize the power needed be it directly [8, 9] optimize or indirectly [10–12]. Another parameter that can affect the drone's ability to do its task is the environment [13, 14].

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 418–428, 2026.
https://doi.org/10.1007/978-981-95-9233-3_30

The environment that the drone was used especially for monitoring can be a limited environment like a recently disaster-affected area or an underdeveloped rural area. In this environment, the connection between the deployed drone and the main controller might receive interference or simply be unavailable. In this situation, decentralized scheduling can be better than the usual centralized method as it only needs communication between drones [15–17]. While for the recently disaster-affected area drone usage is quite obvious as it can be used to assess the affected area [18] as for the rural area it is less so. A possible usage for the rural area is to do a routine monitor for a power generator like a wind turbine [19]. However not just the generator that needs to be inspected, the power line also needs to be inspected to minimize possible damage that can be done to it like an entanglement by a kite [20] which is rarely a point of consideration for monitoring targets. Another consideration due to the lack of available infrastructure is the drone battery as the charging station become far more limited than usual [21–23]. With all the considerations above, a decentralized drone scheduling method concerning drone lifetime is needed. Because of that a decentralized scheduling method using chaos-based bidding consensus was proposed to minimize drone overall travel distance to limit battery usage with the assumption of limited network area.

2 Materials and Method

In this section, we describe the task assignment method we use. With consideration of minimizing the drone travel distance and the limited environment, we use a Chaos-based bidding consensus due to being a decentralized scheduling method while maintaining a fast consensus. We also compare the chaos-based bidding consensus with greedy scheduling to see the difference between the two methods.

2.1 Chaos-Based Bidding Consensus

The chaos-based bidding consensus is a decentralized task assignment where the agent communicates with each other and reach consensus as to who get the task. The consensus was reached by the bid value of each agent for a certain task [24]. The bid value of each agent was sent using a communication protocol called Chaos which provides an efficient way for a many-to-many interaction [25]. Taking advantage of the communication protocol, the agent will share its bid value at a time [26]. The highest bidder gets the target and the other agent must bid for another target. This was done until all target was assigned to an agent. The bid value in our system used the A* algorithm to find the minimum value for the drone travel distance for n maximum amount of nodes with the cost function (J) can be seen in (1) with $l_{(i-1,i)}$ is the distance between a previously

chosen node and the next. For this algorithm, the lowest value is better and no other consideration for bid value, the drone that wins the bid will be the lowest value possible.

$$J = \sum_{i=1}^{n} l_{(i-1,i)} \tag{1}$$

In our system, the agents were allowed to bid up to 3 tasks using the A* algorithm. To minimize information sharing, the queue for bid value was shared by all agents similar to the method proposed by Mahato et al. [27] unlike previous work where every agent shared all of its bids for every target with other agents. Unlike [27] as mentioned above, we let the agent bid a trajectory instead of a single target to minimize time when the agent was forced to take a target that far away. A graphic representation of this method can be seen in Fig. 1.

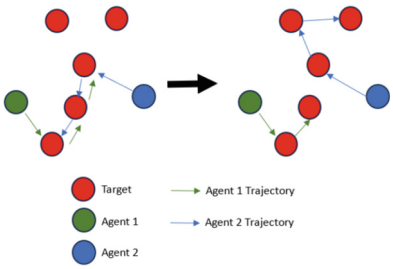


Fig. 1. Representation of bidding consensus with disputed bidding and new consensus.

Considering the above scheduling method characteristics, Algorithm 1 shows the pseudo-code of this method. In this algorithm, we only consider the distance of the last point the agent was at to the target waypoint as the bid. If the waypoint that the agent chooses did not get disputed by another agent then the original agent can take that waypoint. If the waypoint is disputed, the agent may choose another target but it can keep the trajectory before the target to choose another waypoint from there.

Algorithm 1: chaos-based bidding consensus pseudo code

```
Define Waypoint
Define Origin
While there are unassigned waypoints
    For remaining waypoint
        If length drone1 < smallest && length drone1 < waypoint bid
            Smallest drone1 = length drone1
        Else if length drone1 < small1 && length drone1 < waypoint bid
            Small1 drone1 = length drone1
        Else if length drone1 < small2 && length drone1 < waypoint bid
            Small2 drone1 = length drone1
        End if
    End for
    Place smallest drone1 waypoint bid
    Place small1 drone1 waypoint bid
    Place small2 drone1 waypoint bid
    For remaining waypoint
        If length drone2 < smallest && length drone2 < waypoint bid
            Smallest drone2 = length drone2
        Else if length drone2 < small1 && length drone2 < waypoint bid
            Small1 drone2 = length drone2
        Else if length drone2 < small2 && length drone2 < waypoint bid
            Small2 drone2 = length drone2
        End if
    End for
    Place smallest drone2 waypoint bid
    Place small1 drone2 waypoint bid
    Place small2 drone2 waypoint bid
    For remaining waypoint
        If bid drone1 < bid drone2
            Pop waypoint to waypoint drone1
        Else if bid drone1 > bid drone2
            Pop waypoint to waypoint drone2
        End if
    End for
End while
```

2.2 Greedy Scheduling

In this research, we use the Greedy algorithm as a comparison with the Chaos-based bidding consensus. The greedy algorithm is a searching algorithm that finds a global solution through a series of local solution [28] with minimizing the cost function (J) shown in (2). For the scheduling problem, the greedy algorithm was used to give an agent a local solution in this case the smallest distance for the agent unless another agent is closer to the target. If the agent does not get the target, the agent will seek another closer target to fulfill. This action is repeated until all targets are assigned to an agent

[29]. The graphical representation of this method can be seen in Fig. 2.

$$J = l_{(i-1,i)} \tag{2}$$

Fig. 2. Representation of greedy algorithm scheduling with disputed target and new allocation.

The algorithm for the Greedy task scheduling can be seen in algorithm 2. The agent needs to find a local solution with priority given to the smallest distance to the point.

Algorithm 2: greedy algorithm scheduling pseudo code

```
Define Waypoint
Define Origin
While there are waypoints
    For remaining waypoint
        If length drone1 < length drone2
            If length drone1 < smallest drone1
                Smallest drone1 = length drone1
            End if
        Else if length drone1 > length drone2
            If length drone2 < smallest drone2
                Smallest drone1 = length drone1
            End if
        End if
    End for
    Pop smallest drone1 to waypoint drone1
    Pop smallest drone2 to waypoint drone2
End while
```

3 Result

We simulated the Chaos-based bidding consensus to see how it compares with existing method in the aspect of task assignment. The simulation was done with two drones and eight observation tasks. Both drones can deal with all observation tasks equally with each task does not require the drone to hover for a specific amount of time. In this simulation, the drone can only move in a 2D direction. The simulation was done 10 times for each method and presented the most dominant route the method made.

3.1 Task Assignment

In this simulation, we compared the Chaos-based bidding consensus with a greedy method. Both methods require assigning two drones to handle 8 observation targets represented as a waypoint before returning to the origin point. The waypoints themselves were scattered semi-randomly in a 20 by 20-unit area. The coordinate for origin at (0,0). Both drones were scheduled to handle all 8 waypoints while minimizing the total distance traveled for each drone. The result of both simulations can be seen in Fig. 3.

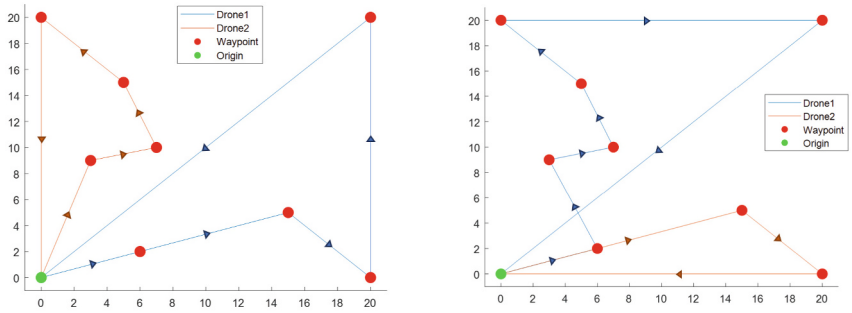


Fig. 3. Result of task scheduling with (a) chaos-based bidding consensus (b) greedy scheduling.

The chaos-based bidding consensus divided the task into two distinct areas that were handled by each drone. This method makes sure that each drone has a balanced number of tasks for each drone. With communication between the two drones, a consensus can be achieved between the two drones to balance the task given to them. For comparison, the following is the result of the greedy method which can be seen in Fig. 3.

The result of the greedy method is quite odd. The greedy method favored Drone 1, giving it priority for task assignment. This favoritism however, resulted in an odd assignment by the system which can be seen in the waypoint at coordinate (6,2) that was assigned to drone 1 when drone 2 can handle it as the drone was assigned to handle the waypoint at the coordinate (15,5). This oddity can be explained by the greedy method's tendency to follow the highest prize or in this case the shortest distance with little consideration to the other drone a contrast with the cooperation that happens at the chaos-based bidding consensus simulation. Table 1 presents a comparison between the chaos-based bidding consensus and the greedy method.

The table shows that the maximum distance travel which is the drone 1 for the chaos-based bidding consensus method was lower by 9.69% than the greedy method. However, for the chaos-based bidding consensus, drone 2 was 7.42% higher than the greedy method. With a lower overall distance, the chaos-based bidding is better than the greedy method for a more time-sensitive case. With the minimum distance travel, it is shown that the balance task assignment approach is done by the chaos-based bidding consensus. The table clearly illustrates this approach by displaying the number of tasks assigned to each drone.

Table 1. Distance travel and task assignment comparison between two methods

Method	Consensus method		Greedy	
	Distance	Task	Distance	Task
Drone 1	71.1667 unit	4 tasks	78.8039 unit	6 tasks
Drone 2	46.0662 unit	4 tasks	42.8824 unit	2 tasks

3.2 Drone Simulation

In this simulation, we implement the assigned task from the previous simulation to simplify drone dynamics with acceleration, as seen in Eqs. (3), (4), and (5) [30]. Where k_t is the thrust drag coefficient, m is drone mass, T is the total thrust, ψ is yaw, θ is pitch, and ϕ is roll. Table 2 displays the comparison of time required by each drone.

$$\ddot{x} = \frac{-1}{m}\left[k_{t_x}\dot{x} + T(\sin\phi\sin\psi + \cos\phi\cos\psi\sin\theta)\right] \tag{3}$$

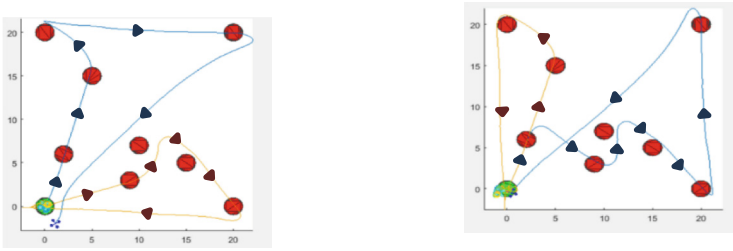
$$\ddot{y} = \frac{-1}{m}\left[k_{t_y}\dot{y} + T(\sin\phi\cos\psi - \cos\phi\sin\psi\sin\theta)\right] \tag{4}$$

$$\ddot{z} = \frac{-1}{m}\left[k_{t_z}z - \dot{m}g + T\cos\phi\cos\theta\right] \tag{5}$$

Table 2. Time comparison between chaos-based bidding consensus and greedy scheduling.

Method	Consensus method	Greedy
Drone 1	48.5 s	58.4 s
Drone 2	32.1 s	27.9 s

Similar to distance travel, the chaos-based bidding consensus takes lower overall time needed than the greedy by 16.95%. The Drone 2, on the other hand, requires 15.05% more time. Fig. 4 shows the drone's trajectory with each scheduling method.



(a) (b)

Fig. 4. Trajectory using (a) chaos-based bidding consensus (b) greedy algorithm scheduling.

4 Discussion

The chaos-based bidding consensus as shown in the simulation can be used for drone scheduling in limited environments with lower travel distances as the drones only need to communicate with each other. The result of the chaos-based bidding consensus being more balanced in its task assignment was not taken into consideration. While this makes the drone's overall travel time lower by 16.95%, it also makes the maximum tasks handled by the drone decrease by 50%. In return, the drone requires less observation data. This method minimizes the resources required for each drone to perform its task, taking into account the network limitation. While this method has benefits, it also has limitations. Communication range can pose a significant challenge in a network environment with limited resources. Due to the on-the-fly bidding process, if two or more drones bid for the same task but are outside the communication range, they may end up bidding for the same task. If we had used the centralized method for scheduling or reached a consensus before the drone's action, we could have prevented this issue. However, this approach undermined the adaptability of the decentralized method, which responds to unforeseen circumstances during operation. Once communication is reestablished, this problem can be resolved, but it results in wasted time attempting to complete a task that someone else can perform more efficiently.

5 Conclusion

The chaos-based bidding consensus method shows a balanced task scheduling, which results in a lower time needed and distance travel overall. Compared to the greedy method, the chaos-based bidding method requires 16.95% less time and 9.69% less distance for the highest requirement. This approach also comes with fewer overloading tasks for a single agent making it a suitable choice for a resource-critical environment. The chaos-based bidding method can effectively schedule drone tasks, especially in networks with few resources and resource-critical environments like disaster-affected areas or rural areas that aren't well developed.

Future studies in this area may consider a few potential constraints for the simulation. As previously discussed, the first limitation is the communication range limitation, which can result in time wastage. Another factor to consider is network latency. Although

chaos-based communication partially addresses it, drone completion time may need to see the impact of latency. In addition to the previously mentioned limitations, a hardware implementation of this method could be conducted to assess its resilience to unexpected incidents.

Acknowledgement. This research is funded by the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia under the KATALIS project with contract number 059/E5/PG.02.00/PL.BATCH.2/2024, 016/SP2H/PL.BATCH.2/LL4/2024, and 152/LIT07/PPM-LIT/2024

References

1. Pasha J et al (2022) The drone scheduling problem: a systematic state-of-the-art review. IEEE Trans Intell Transp Syst 23:14224–14247. https://doi.org/10.1109/TITS.2022.3155072
2. Chung SH, Sah B, Lee J (2020) Optimization for drone and drone-truck combined operations: a review of the state of the art and future directions. Comput Oper Res 123:105004. https://doi.org/10.1016/j.cor.2020.105004
3. Agatz N, Bouman P, Schmidt M (2018) Optimization approaches for the traveling salesman problem with drone. Transp Sci 52:965–981. https://doi.org/10.1287/trsc.2017.0791
4. Torabbeigi M, Lim GJ, Kim SJ (2020) Drone delivery scheduling optimization considering payload-induced battery consumption rates. J Intell Robot Syst 97:471–487. https://doi.org/10.1007/s10846-019-01034-w
5. Qi H, Hu Z, Huang H, Wen X, Lu Z (2020) Energy efficient 3-D UAV control for persistent communication service and fairness: a deep reinforcement learning approach. IEEE Access 8:53172–53184. https://doi.org/10.1109/ACCESS.2020.2981403
6. Kim J, Song BD, Morrison JR (2013) On the scheduling of systems of UAVs and fuel service stations for long-term mission fulfillment. J Intell Robot Syst 70:347–359. https://doi.org/10.1007/s10846-012-9727-0
7. Song BD, Kim J, Kim J, Park H, Morrison JR, Shim DH (2013) Persistent UAV service: an improved scheduling formulation and prototypes of system components. In: 2013 international conference on unmanned aircraft systems (ICUAS). IEEE, pp 915–925. https://doi.org/10.1109/ICUAS.2013.6564777
8. Zhang J, Wu Y, Min G, Hao F, Cui L (2020) Balancing energy consumption and reputation gain of UAV scheduling in edge computing. IEEE Trans Cogn Commun Netw 6:1204–1217. https://doi.org/10.1109/TCCN.2020.3004592
9. Jung S, Kim J, Kim J-H (2020) Joint message-passing and convex optimization framework for energy-efficient surveillance UAV scheduling. Electron (Basel) 9:1475. https://doi.org/10.3390/electronics9091475
10. Hu M et al (2019) On the joint design of routing and scheduling for vehicle-assisted multi-UAV inspection. Futur Gener Comput Syst 94:214–223. https://doi.org/10.1016/j.future.2018.11.024
11. Hu M et al (2019) Joint routing and scheduling for vehicle-assisted multidrone surveillance. IEEE Internet Things J 6:1781–1790. https://doi.org/10.1109/JIOT.2018.2878602
12. Weinstein A, Schumacher C (2007) UAV scheduling via the vehicle routing problem with time windows. In: AIAA Infotech@aerospace 2007 conference and exhibit. American Institute of Aeronautics and Astronautics, Reston, Virigina. https://doi.org/10.2514/6.2007-2839
13. Zhao J, Liu J, Jiang J, Gao F (2020) Efficient deployment with geometric analysis for mmWave UAV communications. IEEE Wirel Commun Lett 1–1. https://doi.org/10.1109/LWC.2020.2982637

14. Song BD, Kim J, Morrison JR (2014) Towards real time scheduling for persistent UAV service: a rolling horizon MILP approach, RHTA and the STAH heuristic. In: 2014 international conference on unmanned aircraft systems (ICUAS). IEEE, pp 506–515. https://doi.org/10.1109/ICUAS.2014.6842292
15. Hassija V, Saxena V, Chamola V (2020) Scheduling drone charging for multi-drone network based on consensus time-stamp and game theory. Comput Commun 149:51–61. https://doi.org/10.1016/j.comcom.2019.09.021
16. Ghazzai H, Kadri A, Ben Ghorbel M, Menouar H, Massoud Y (2019) A generic spatiotemporal UAV scheduling framework for multi-event applications. IEEE Access 7:215–229. https://doi.org/10.1109/ACCESS.2018.2885321
17. Cummings ML, Clare A, Hart C (2010) The role of human-automation consensus in multiple unmanned vehicle scheduling. Hum Factors: J Hum Factors Ergon Soc 52:17–27. https://doi.org/10.1177/0018720810368674
18. Ejaz W, Ahmed A, Mushtaq A, Ibnkahla M (2020) Energy-efficient task scheduling and physiological assessment in disaster management using UAV-assisted networks. Comput Commun 155:150–157. https://doi.org/10.1016/j.comcom.2020.03.019
19. Cao P, Liu Y, Qiu M, Yang C, Xie S (2019) MEC-driven UAV routine inspection system in wind farm under wind influence. In: 2019 12th international conference on intelligent computation technology and automation (ICICTA). IEEE, pp 672–677. https://doi.org/10.1109/ICICTA49267.2019.00148
20. Sasmono S et al (2022) The resilience of the grid from the risk of failure due to kite thread disturbance: case study Khatulistiwa grid. In: 2022 IEEE PES innovative smart grid technologies—Asia (ISGT Asia). IEEE, pp 120–124. https://doi.org/10.1109/ISGTAsia54193.2022.10003494
21. Li J, Cao X, Guo D, Xie J, Chen H (2020) Task scheduling with UAV-assisted vehicular cloud for road detection in highway scenario. IEEE Internet Things J 7:7702–7713. https://doi.org/10.1109/JIOT.2020.2992088
22. Duan T, Wang W, Wang T, Chen X, Li X (2020) Dynamic tasks scheduling model of UAV cluster based on flexible network architecture. IEEE Access. 8:115448–115460. https://doi.org/10.1109/ACCESS.2020.3002594
23. Macrina G, Di Puglia Pugliese L, Guerriero F, Laporte G (2020) Drone-aided routing: a literature review. Transp Res Part C Emerg Technol 120:102762. https://doi.org/10.1016/j.trc.2020.102762
24. Sarkar C, Dey S, Agarwal M (2018) Semantic knowledge driven utility calculation towards efficient multi-robot task allocation. In: 2018 IEEE 14th international conference on automation science and engineering (CASE). IEEE, pp 144–147. https://doi.org/10.1109/COASE.2018.8560555
25. Landsiedel O, Ferrari F, Zimmerling M (2013) Chaos. In: Proceedings of the 11th ACM conference on embedded networked sensor systems. ACM, New York, NY, USA, pp 1–14. https://doi.org/10.1145/2517351.2517358
26. Turner J, Meng Q, Schaefer G, Soltoggio A (2018) Fast consensus for fully distributed multi-agent task allocation. In: Proceedings of the 33rd annual acm symposium on applied computing. ACM, New York, NY, USA, pp. 832–839. https://doi.org/10.1145/3167132.3167224
27. Mahato P, Saha S, Sarkar C, Shaghil M (2023) Consensus-based fast and energy-efficient multi-robot task allocation. Robot Auton Syst 159:104270. https://doi.org/10.1016/j.robot.2022.104270
28. Wang C, Ge P, Sun L, Wang F (2022) Research on user-side flexible load scheduling method based on greedy algorithm. Energy Rep 8:192–201. https://doi.org/10.1016/j.egyr.2022.10.352

29. Zhao Z, Zhou M, Liu S (2022) Iterated greedy algorithms for flow-shop scheduling problems: a tutorial. IEEE Trans Autom Sci Eng 19:1941–1959. https://doi.org/10.1109/TASE.2021.3062994
30. Islam M, Okasha M, Idres MM (2017) Dynamics and control of quadcopter using linear model predictive control approach. IOP Conf Ser Mater Sci Eng 270:012007. https://doi.org/10.1088/1757-899X/270/1/012007

Development of Pressure Control and Oxygen Control Methods for Continuous Mandatory Ventilation (CMV) and Synchronized Intermittent Mandatory Ventilation (SIMV) Modes in Ventilator

Basari[1,2(✉)] and Sulthon Abdul Aziz[1]

[1] Biomedical Engineering Program, Department of Electrical Engineering, Faculty of Engineering, University of Indonesia, Depok, Indonesia
basari.st@ui.ac.id

[2] Faculty of Engineering, Research Center for Biomedical Engineering, University of Indonesia, Depok, Indonesia

Abstract. Research on ventilators is increasing along with the increasing human need for mechanical breathing aids. A mechanical ventilator is a device that can provide respiratory support to maintain life in patients who require long-term ventilation assistance due to respiratory disorders or failure. The development of pressure control is vital to avoid barotrauma or hyperventilation, and the development of oxygen control is essential to prevent hypoxia or hyperoxia. This study uses a microcontroller with better performance than previous studies and complements the needs of health workers in providing medical interventions to patients. This research aims to develop a pressure control and oxygen control method using STM32F407VGT6 as a microcontroller with a fuzzy logic method as a closed-loop control system. The results of this development showed that the accuracy of the inspiratory pressure parameter in pressure control was tested with a VT analyzer, and the accuracy of the output results of the tool was tested at 91.98% to 98.54%. In comparison, oxygen control had an accuracy of 89.9% to 99.6%. These results of the performance test comply to the BPAFK standards, Ministry of Health Republic of Indonesia for maximum error in accuracy.

Keywords: FiO_2 · Oxygen Control · Pressure Control · Pressure Inspiration · Ventilator

1 Introduction

Research on ventilators is increasing along with the increasing human need for mechanical breathing aids. The COVID-19 pandemic is one of the leading causes of death worldwide. It is estimated that 15–20% of COVID-19 sufferers require hospitalization, and 3–5% require treatment in the Intensive Care Unit (ICU). One of the contributing factors is the lack of a volumetric fraction of oxygen in the inhaled gas, known as FiO_2 [1].

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 429–445, 2026.
https://doi.org/10.1007/978-981-95-9233-3_31

In High-Flow Nasal Cannula (HFNC) therapy, a stable and appropriate FiO_2 is needed for treatment, especially in infant and neonatal patients [2] and COVID-19 patients [3]. Accurate inspired oxygen fraction is also essential to determining Non-Invasive Positive Pressure Ventilation (NIPPV) and Continuous Positive Airway Pressure (CPAP) therapy for patients with Obstructive Sleep Apnea Syndrome (OSAS) who experience acute respiratory failure [4, 5]. Efficient and precise oxygenase is an important factor in supporting patient recovery. Excessive fractions of inspired oxygen can cause patient poisoning, so the mixing process must be maintained appropriately. Several methods have been reported in the literature. Closed-loop control systems have been implemented in several studies using various algorithms, such as PID, Fuzzy, and neuro-fuzzy [6–9].

A mechanical ventilator is a device that can provide respiratory support to maintain life in patients who require long-term ventilation assistance due to respiratory disorders or failure [10]. Various ventilators are used, including pressure control mode, volume control mode, dual control mode, assist control mode, etc. [11]. Making a stable mechanical ventilator is tricky because parameters constantly change during breathing, namely compliance [12]. Therefore, a control system is needed to adapt to parameter changes. Changes in compliance influence the achievement of the desired inspiratory pressure and inspiratory flow [13].

In the pressure control method, several studies in the last five years have used Arduino and Teensy 4.0 as microcontrollers, the hardware used also varies to different control system methods. The use of the STM32F407VGT6 microcontroller in the pressure control system is based on the advantages of better performance, many I/O pins so that it can support the needs of a complex ventilator system compared to microcontrollers in previous studies. The NI myRIO Embedded Device (myRIO-1900) is superior but very expensive so that the use of the STM32F407VGT6 can reduce material costs. The use of solenoid valves, stepper motors, and proportional valve settings are also used to improve accuracy in producing FiO_2 levels compared to DC motors and unidirectional valves, and to avoid the factor of decreasing material quality in ambu bags over time. Previous research has Steady State Error (SSE) values of 0.2 to 4.5% and Mean Absolute Percentage Error (MAPE) 1.01%, while in this study STM32F407VGT6 was used with Fuzzy Logic control system and MAPE was obtained at 3.7%. The novelty of this pressure control research is shown in Table 1.

The novelty of Oxygen Control research is shown in Table 2 that many studies in the last five years have used Arduino as a microcontroller, and the methods used by researchers vary. The use of the STM32F407VGT6 microcontroller in the oxygen control system is based on the advantages of better performance, and many I/O pins so that it can support the needs of a complex ventilator system compared to microcontrollers in previous studies. The use of solenoid valves, stepper motors, and proportional control valves are also used to improve accuracy in producing FiO_2 levels and avoid factors that reduce the quality of materials in ambu bags over time. The error values obtained from previous studies vary, ranging from Steady State Error (SSE) 1.15% to 3.29%, in this study, STM32F407VGT6 was used as a microcontroller to carry out oxygen control tasks using the pneumatic valve and stepper motor method with a Mean Absolute Percentage Error (MAPE) of 6.075%.

Table 1. State of the art pressure control research

Microcontroller	Methods	Control system	Error value	References
Teensy 4.0	Proportional control valve	PI control	SSE 1.25%	[14]
Arduino UNO	Solenoid valve and unidirectional valve	Fuzzy logic	MAPE 1.01%	[15]
NI myRIO embedded device (myRIO-1900)	BVM (bag valve mask)	PI and LQR control	SSE 4.5% and 0.2%	[16]
Arduino Nano	DC geared motor and ambu bag	n.a.	n.a.	[17]
Arduino UNO and ESP32 microcontroller	BVM (bag valve mask)	n.a.	n.a.	[18]
STM32F407VGT6	Solenoid valve, TMC2208 stepper motor, and proportional control valve	Fuzzy logic	MAPE 3.7%	This research

Table 2. State of the art oxygen control research

Microcontroller	Methods	Error value	References
ATMega328P	Pneumatic valve	SSE 3.29%	[9]
Raspberry Pi Pico	High pressure turbine and pneumatic valve	$< 1 cmH_2O$	[19]
Arduino UNO	Ambu bag and TB6600 stepper motor	n.a.	[20]
Teensy 4.0	Proportional control valve	SSE 1.15%	[1]
Microcontroller with PWM control	DC motor conrolled	n.a.	[21]
STM32F407VGT6	Solenoid valve, TMC2208 stepper motor, and proportional control valve	MAPE 6.075%	This research

Our previous developments used the Arduino® Mega 2560 microcontroller with a clock speed of 16 MHz and 8 KB SRAM, and available in two modes, namely Continuous Mandatory Ventilation (CMV) mode and Continuous Positive Airway Pressure (CPAP) mode. The user cannot adjust the fraction of inspired oxygen (FiO_2) value in the previous version, as the ventilator's venturi design determines the FiO_2 value, which remains fixed.

The controls on previous versions of ventilators used an open-loop control system for volume and alarm feedback if the system detected pressure exceeding the maximum [22].

This research is an advanced development of the Covent-20 earlier version. This research uses STM32F407VGT6 with the addition of modes to VC-CMV (Volume Controlled - Continuous Mandatory Ventilation), PC-CMV (Pressure Controlled - Continuous Mandatory Ventilation), VC-SIMV (Volume Controlled - Synchronous Intermittent Mandatory Ventilation), PC-SIMV (Pressure Controlled - Synchronous Intermittent Mandatory Ventilation), and CPAP (Continuous Positive Airway Pressure), where in all these modes a button has also been added for the user to adjust the FiO_2 to be given to the patient to perform pressure control and oxygen control in CMV and SIMV modes. The inspiratory pressure and FiO_2 parameters will be tested for validity using the VT650 Gas Flow Analyzer from Fluke Biomedical™ (called VT Analyzer) by comparing the user set values, the tool output values, and the VT Analyzer reading results. The results of this validity test are also compared with the Health Equipment and Facility Security Office (BPAFK), Ministry of Health Republic of Indonesia to meet the national standards. This development has used a microcontroller with better specifications but has not been integrated with more advanced technologies such as Artificial Intelligence (AI) and Internet of Things (IoT) but can reach the basic functional needs of the ventilator.

2 Materials and Method

2.1 Research Stage

As an initial research stage, a literature review was carried out to look for references regarding pressure control modes and oxygen mixing methods in ventilators, design plans, and measurement and testing methods for ventilators. From the literature study results, the control system, electrical system, and mechanical system were designed in parallel because these three designs have a relationship with each other, so the design process must be carried out in parallel. After designing the three systems, integration is carried out to test the success of the three designed systems and check whether or not conflicting systems cause conflicts. The integrated system was tested using a measuring instrument in the form of a ventilator gas flow analyzer [23]. The ventilator gas flow analyzer will validate all parameters and outputs produced by the system. When measurement results differ, evaluation and changes are made to the design of the three systems according to the source of the problem of the differences encountered. The test results of the system created are then reported to serve as material for further evaluation and development. The flow chart of the stages of this research can be seen in Fig. 1.

Our earlier running research used STM32F411CEU6 with a maximum clock speed of 100 MHz and 48-pin UFQFPN [24] while the development was in research. It uses an STM32F407VGT6 with a maximum clock speed of 168 MHz with LQFP 100 pins [25]. The replacement of this microcontroller is based on increasing the electronic components used so that the STM32F411CEU6 can no longer accommodate the number of pins needed. Apart from that, the STM32F407VGT6 also has a higher maximum clock speed to increase system performance.

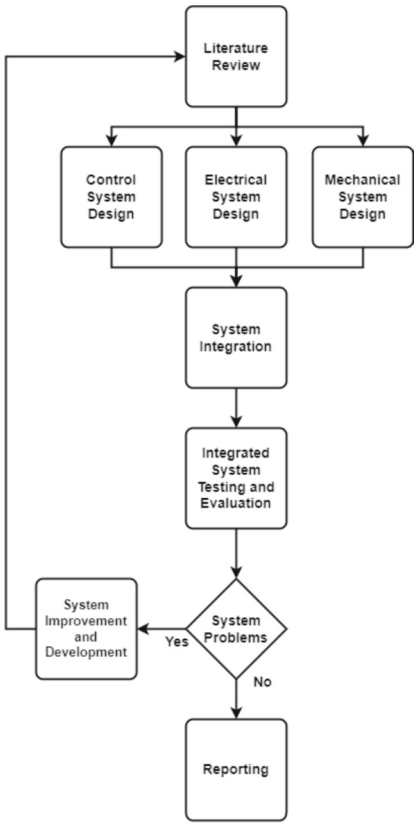


Fig. 1. Research stages flow chart

Developments in this research include the addition of a stepper motor to create a system that can regulate FiO_2 levels according to the operator's needs to carry out medical interventions on patients. In addition, to support the system for regulating FiO_2 levels, the system requires an additional sensor in the form of an oxygen sensor so that the system can check the commands carried out. The addition of components was also carried out considering the VT_e (Tidal Expiration Volume) parameter, which is one of the standard testing parameters for the Health Equipment and Facility Security Office (BPAFK), Ministry of Health Republic of Indonesia [26], so an additional air flow sensor is needed to measure the air volume during expiration. In addition, in this study, the Positive End Expiratory Pressure (PEEP) value, which was previously determined by hardware components that were manually adjusted by hand, was changed to be automatic based on the input provided. To support this PEEP setting, a solenoid valve must be added to the expiratory airflow to close when the expected PEEP value has been reached.

2.2 Pressure Control Design

When using the pressure control mode on the ventilator, the operator does not directly regulate the tidal volume and airflow but rather regulates the pressure applied to the lungs over a period so that the lungs expand to a certain volume [27]. The same rules apply, such as calculating the IBW (Ideal Body Weight) and tidal volume. The operator will still measure the patient's height and calculate the IBW and the appropriate tidal volume range. However, instead of directly regulating the tidal volume, the operator will regulate the applied pressure and then monitor the lung volume after about three breaths [28]. If the volume generated is too high, decrease the pressure control (PC) by 2 cmH_2O. If the volume is too low, increase the PC by 2 cmH_2O and evaluate. As a safe starting point, pressure control is recommended to start at 14 cmH_2O. With a slightly higher PEEP (Positive End-Expiratory Pressure), this will ensure that the peak pressure remains well below 35 cmH_2O [29].

The pressure control flow chart is shown in Fig. 2. This flow chart has been explained previously but concisely focuses on pressure control. This flowchart explains that the pressure control mode starts with the user setting the inspiratory pressure value in PC-CMV mode and PC-SIMV mode [29, 30]. The stepper motor moves to the specified position according to the data array that matches the other settings. The pressure data obtained from the MPX5010DP as a pressure sensor is processed from the output voltage received by the STM32F407VGT6 microcontroller. Pressure is calculated from Eq. (1).

$$P = \frac{Vout - error}{s - offset} \tag{1}$$

where P represents the pressure measured (in cmH_2O), Vout denotes the voltage output from the sensor, *error* signifies the coefficient of pressure error influenced by temperature factors, s defines the coefficient of the pressure sensor's sensitivity, and *offset* corresponds to the deviation value derived from the measurement.

Where the pressure value is taken from the read voltage data which is reduced by the coefficient for the error factor which is divided by the sensor sensitivity coefficient and the sensor offset coefficient. Pressure data is sampled at a frequency of 100 kHz, but the speed of data display has a limit so that it cannot display all pressure data that has been read. The transfer function in this pressure control system is expressed in Eq. (2) and (3)

$$y_p(t) = G_p(s) \cdot u_p(t) \tag{2}$$

$$G_p(s) = \frac{K_p}{\tau_p s + 1} \tag{3}$$

where $y_p(t)$ is the actual pressure of the system, $G_p(s)$ epresents the physical system model (stepper control), $u_p(t)$ denotes the pressure control signals generated by the fuzzy controller, K_p s the gain of the stepper for pressure control, and $\tau_p s$ refers to the time constant of the stepper for pressure control.

The solenoid valve will open/close according to other parameter settings besides inspiratory pressure, and then the resulting parameters are read. If the inspiratory pressure

reading is appropriate, then control is complete, and control will be carried out for the next breath. If the inspiratory pressure is not appropriate, evaluation and correction are carried out using fuzzy control on the data array to change the step to improve the resulting inspiratory pressure value.

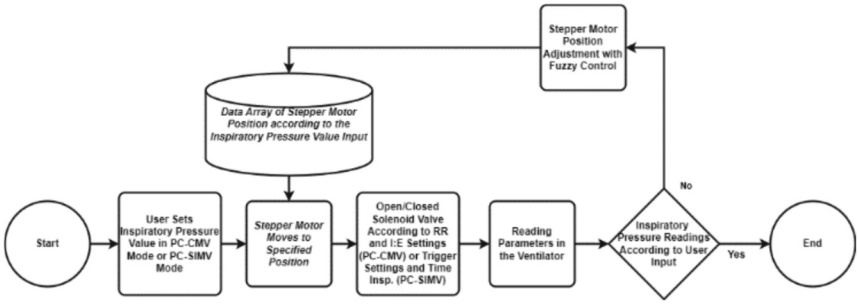


Fig. 2. Pressure control flow chart

2.3 Oxygen Control Design

Oxygen on a mechanical ventilator is always regulated, regardless of the method or type of breathing used. Because the air delivered by a mechanical ventilator can affect the overall air the patient breathes, the composition or percentage of oxygen must be adjusted on the ventilator. Oxygen regulation on a ventilator is done through FiO_2. The operator will adjust the FiO_2 between 0.21 and 1.0 (21% to 100%) to ensure the patient receives enough oxygen [31]. Oxygen control is expressed in the Fraction of Inspired Oxygen (FiO_2) value. FiO_2 is the percentage of oxygen in the air mixture delivered by the ventilator during each breathing cycle [32]. Increasing FiO_2 will increase the patient's oxygen saturation [33].

The oxygen control flow chart is shown in Fig. 3. This flow chart has been explained previously but is kept concise and focuses on oxygen control. In this flow chart, it is explained that the oxygen control mode starts with the user setting the inspiratory pressure value in all available modes, then stepper motors 1 and 2 moves to the specified position according to the data array that has been created according to the desired oxygen gas mixing percentage value [34, 35]. The oxygen mixing was obtained from calculating the airflow in liters per minute (LPM) and oxygen flow in liters per minute by using Eq. (4) and Eq. (5) to produce the oxygen fraction (FiO_2) at the appropriate inspiratory flow [36].

$$Q_{air} = \frac{1-x}{0.79} QT \tag{4}$$

$$Q_{O2} = \frac{x-0.21}{0.79} QT \tag{5}$$

Eq. (4) sets the setpoint for the air flow regulator (Q_{air}), while Eq. (5) sets the setpoint for the oxygen flow regulator (Q_{O2}). This ratio determines the amount of oxygen content mixed in the mixing tank. The total flow QT is the accumulation of Q_{O2} and Q_{air} [5]. The transfer function in this pressure control system is expressed in the Eqs. (6), (7), and (8).

$$y_m(t) = G_o(s) \cdot G_a(s) \cdot u_m(t) \tag{6}$$

$$G_o(s) = \frac{K_o}{\tau_o s + 1} \tag{7}$$

$$G_a(s) = \frac{K_a}{\tau_a s + 1} \tag{8}$$

where $y_m(t)$ is the actual mixed oxygen of the system, $G_o(s)$ represents the physical system model for the stepper control of the oxygen source, $G_a(s)$ denotes the physical system model for the stepper control of the ambient source, $u_m(t)$ is the mixed oxygen control signal generated by the fuzzy controller, K_o is the gain of the stepper control for the oxygen source, K_a is the gain of the stepper control for the ambient source, $\tau_o s$ represents the time constant of the stepper control for the oxygen source, and $\tau_a s$ corresponds to the time constant of the stepper control for the ambient source.

The solenoid valve will open/close according to other parameter settings in each mode, and then the resulting parameters are read. This oxygen control system is still made with an open loop control system, so there are no corrections or evaluations in this flow chart.

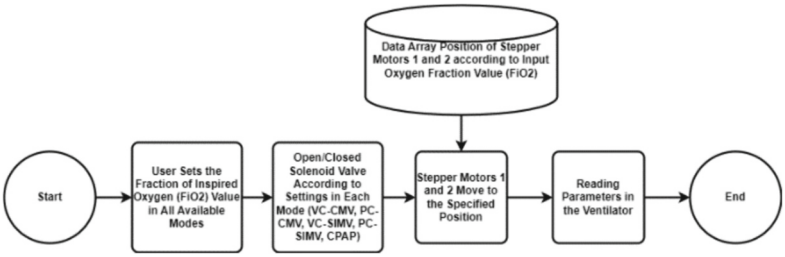


Fig. 3. Oxygen control flow chart

2.4 Standardized Criteria Calculation

In this study, data processing uses a formula to calculate the accuracy of each measurement to produce an output of parameters that comply with standard limits. This formula is used to analyze pressure control and oxygen control testing of 100 data from each parameter. The data retrieval process is through averaging 100 retrieved points to be used as data to represent valid data. Data retrieval is done by taking measurements on each respiratory cycle by comparing the data obtained from the readings on the ventilator to

the readings made with the VT Analyzer, both in the pressure control system and the oxygen control system. Data collection is conducted directly by connecting the ventilator, VT Analyzer, and least lung in series. The test lung used is Selftestlung™ Dräger which has a maximum volume of 1000 mL, resistance of 20 mbar/L/s and compliance of 25 ml/mbar [37]. Data processing is done by the Eq. (9):

3 Average on Measurement

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{9}$$

where $\bar{x}$ is the average measurement value, x_i represents the individual random values of the measurement data ($_1$, $_2$, …, $_I$), and n denotes the total amount of data collected, which influences the error value.

Based on Eq. (9) above, we can calculate the average of ungrouped data. For example, if $_1$, $_2$, and so on are data collected from a sample, they can be symbolized by $\bar{x}$. From the data that has been obtained, it will be entered into a formula so that we get the calculated average [38];

4 Error Value

$$Error = \bar{x} - uut \tag{10}$$

where $\bar{x}$ is the average measurement value, and *uut* represents the standard value or the setting value.

Based on Eq. (10), we can calculate the error value, where we have to find $\bar{x}$ or the average value, which will later be subtracted from the setting value or value set;

5 Error Percentage Value

$$\%Error = \frac{Error}{Setting} \times 100\% \tag{11}$$

where *Error* represents the error value, and *Setting* denotes the value that has been set.

Based on Eq. (11), we can calculate what percentage error value will be found in the tool being made. Namely by looking for the error value and dividing it by the setting value, then multiplying it by 100% [39];

6 Accuracy Percentage Value

$$\%Accuracy = 100\% - \%Error \tag{12}$$

Based on Eq. (12), we can determine the percentage error value obtained in the accuracy.

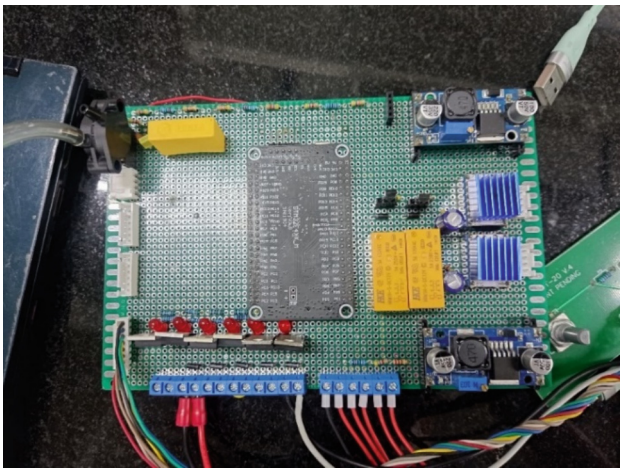


Fig. 4. Ventilator development prototype

7 Result

Design and development of pressure control and oxygen control methods for CMV and SIMV modes on ventilators have been carried out. This development includes adding control for the pressure delivered to the patient. This control system takes the inspiratory pressure value from the PIP value minus the PEEP value. PIP is the highest-pressure limit during inspiration, and PEEP is the final limit of positive pressure at the end of expiration. FiO_2 levels are controlled using two stepper motors connected to a proportional valve to regulate two gas sources: an oxygen gas source and a free air source. The development prototype can be seen in Fig. 4.

7.1 Pressure Control Testing

The VT analyzer used does not display the Pi (Pressure Inspiration) parameter, so the author tries to carry out measurements using the same method, namely by subtracting the PIP value from the PEEP value displayed on the VT analyzer. The set inspiratory pressure values include 8, 12, 14, 16, and 20 cmH_2O. Table 3 summarizes the pressure control method test results.

Table 3 shows that the output accuracy of the pressure control is quite good. This means that the development of the pressure control method has succeeded in fulfilling the command values set by the user with an accuracy of 91.98% to 98.54%. However, Table 3 also shows the high maximum pressure read by the tool. These pressure spikes can be seen more clearly in Fig. 5 and Fig. 6.

Figure 5 shows the pressure control test graph at a setting of 8 cmH_2O. From the 100 data taken, it was seen that the tool read many pressure spikes, indicating instability in the pressure reading process, both fundamentally to the method of reading the PIP and PEEP values , which seemed different from that carried out by the VT analyzer. Fig. 5

Table 3. Summary of pressure control testing

Parameter	Minimum concentration (% FiO_2)	Maximum concentration (% FiO_2)	Average concentration (% FiO_2)	Error value (% FiO_2)	% Error value	% Accuracy
8 cmH_2O						
Machine output	6	11	8.313	0.313	4	96.09
VT analyzer test	6.7	8.1	7.358	0.642	8	91.98
12 cmH_2O						
Machine output	10	15	12.175	0.175	1	98.54
VT analyzer test	12	12.42	12.176	0.176	1	98.3
14 cmH_2O						
Machine output	12	18	14.513	0.513	4	96.34
VT analyzer test	14	14.52	14.21	0.21	2	96.34
16 cmH_2O						
Machine output	14	20	16.692	0.692	4	95.68
VT analyzer test	16.19	16.92	16.493	0.493	3	95.68
20 cmH_2O						
Machine output	18	24	21.13	1.13	6	94.35
VT analyzer test	20.48	21.01	20.817	0.817	4	94.35

shows that the inspiratory pressure readings taken by the VT analyzer fluctuate slightly, so a closed-loop control system is needed for this pressure control method.

Figure 6 shows the pressure control test graph at a setting of 20 cmH_2O. Of the 100 data taken, the system read relatively fewer pressure spikes than at the eight cmH_2O settings, and the inspiratory pressure readings by the VT analyzer were also relatively more stable. Based on these observations, further developments are needed to improve the stability of inspiratory pressure readings and better control programs to produce better accuracy and stability.

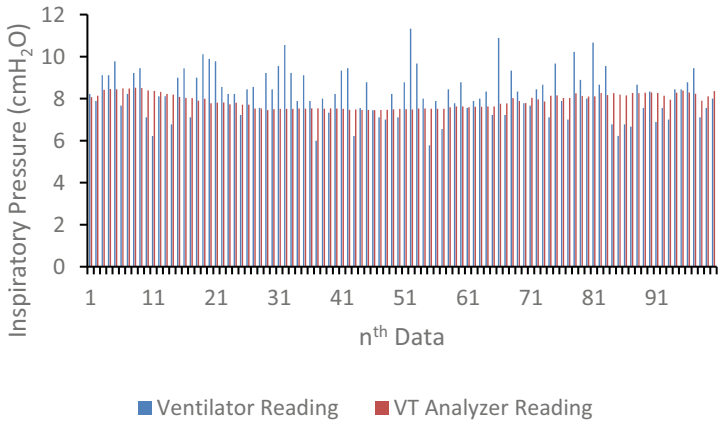


Fig. 5. Inspiratory pressure test at setting 8 cmH_2O

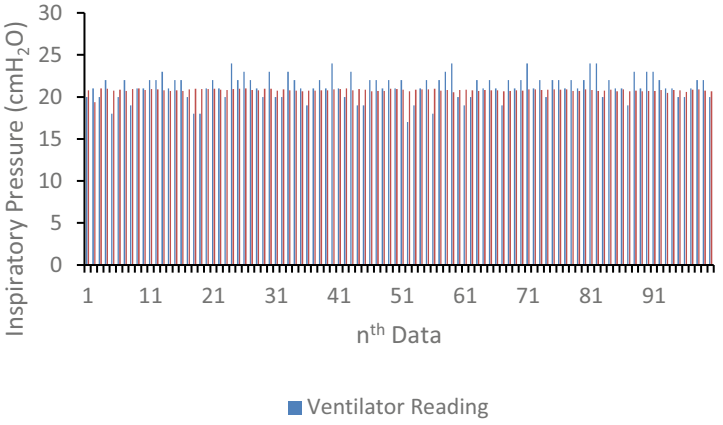


Fig. 6. Inspiratory pressure test at setting 20 cmH_2O

7.2 Oxygen Control Testing

The development of oxygen control is carried out by commanding the stepper motor to move the proportional or airflow control valve. Two airflow control valves regulate two different air sources: the oxygen gas source and the free air source. The oxygen level given to the patient is regulated by adjusting the proportion of openings for the two sources so that the desired oxygen level is achieved.

Figure 7 displays a graph of oxygen control readings in setting FiO_2 40%, 60%, 80%, and 100% changed continuously. Oxygen control is still difficult because the OCS-3F sensor read by the microcontroller gives fluctuating values.

Table 4 shows that the output accuracy of FiO_2 levels is quite good, but it is not good enough, namely 89.9% to 99.6%. This is because the control system still uses an open-loop system, and the OCS-3F oxygen sensor readings are not optimal, resulting in an error in the designed closed-loop control system. Improvements to the OCS-3F sensor

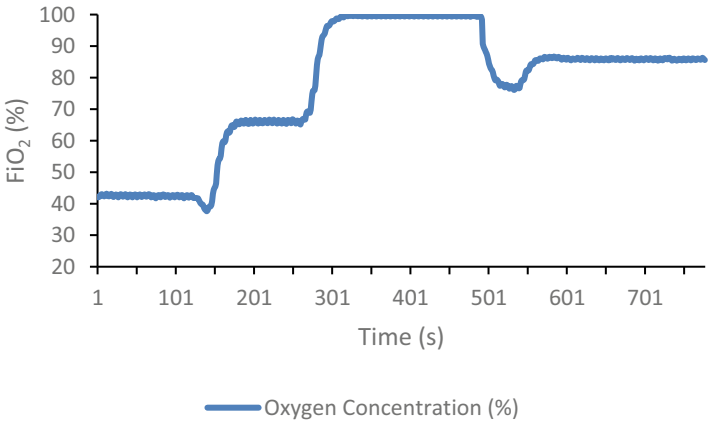


Fig. 7. FiO_2 control reading

readings or replacement of the sensor type need to be carried out so that the closed-loop control system can be implemented to improve oxygen control performance.

Table 4. Testing FiO_2 levels with VT analyzer

Parameter	Minimum concentration (% FiO_2)	Maximum concentration (% FiO_2)	Average concentration (% FiO_2)	Error value (% FiO_2)	% Error value	% Accuracy
Setting 40% FiO_2	42	43.2	42.587	2.58	6.5	93.5
Setting 60% FiO_2	65.3	66.8	66.067	6.067	10.1	89.9
Setting 80% FiO_2	85.4	86.3	85.823	5.823	7.3	92.7
Setting 100% FiO_2	99.4	99.8	99.608	0.392	0.4	99.6

8 Discussions

From the 100 data points taken on each parameter, the pressure spike was read by the system, so further development is needed to improve the stability of the reading on inspiratory pressure and a better control program to produce better accuracy and stability. The output accuracy of the FiO_2 level is not good enough, namely 89.9% to 99.6%. This is because the control system used is still using an open loop system; the reading of the OCS-3F oxygen sensor, which is not optimal, causes the closed loop control system that is designed to be an error. Repairing the OCS-3F sensor reading or replacing the type of sensor needs to be done so that the closed-loop control system can be used to improve oxygen control performance.

In the pressure control method, several studies in the last five years have used Arduino and Teensy 4.0 as microcontrollers, the hardware used also varies to different control system methods with Steady State Error (SSE) values of 0.2 to 4.5% and Mean Absolute Percentage Error (MAPE) 1.01%, while in this study STM32F407VGT6 was used with a Fuzzy Logic control system and a MAPE of 3.7% was obtained [1, 9, 19–21]. The novelty of Oxygen Control research is that many studies in the last five years have used Arduino as a microcontroller, and the methods used by researchers vary. The error values obtained from previous studies vary, ranging from Steady State Error (SSE) 1.15% to 3.29%. In this study, STM32F407VGT6 was used as a microcontroller to carry out oxygen control tasks using pneumatic valves and stepper motors with a MAPE of 6.075% [14–18].

This research is compared with products that have been commercialized in the world, such as Drager, GE, and Hamilton [40–42]. The specifications for oxygen control are the same as those products, but for pressure control, the current range of pressure results are still far compared to those products.

In this study, the average of the instrument indication value was compared with the average of the standard indication value (VT Analyzer), which was then calculated for the correction of both the instrument indication value and the instrument setting. This correction percentage should not exceed 15% to pass the passing limit of each parameter based on the BPAFK standard, Ministry of Health Republic of Indonesia. Based on the results obtained, namely the largest percentage error for pressure control is 8% and 10.1% for oxygen control, the results of this study can meet the required limit of testing based on the BPAFK standard, Ministry of Health Republic of Indonesia.

9 Conclusion

In conclusion, the research successfully developed and tested a pressure and oxygen control method using the STM32F407VGT6, with the inspiratory pressure parameter demonstrating good accuracy (91.98% to 98.54%), although further improvement is needed due to some fluctuations in pressure values. FiO_2 control was also implemented, but the large error margin (accuracy of 89.9% to 99.6%) suggests the need for a closed-loop control system with accurate FiO_2 readings. The results of the accuracy test satisfy the requirement limits of the BPAFK standards, Ministry of Health Republic of Indonesia, which must lower than 15% maximum error in accuracy. Future research will focus on enhancing ventilator performance by using FiO_2 sensor feedback to implement a closed-loop control system and creating more data arrays to improve system responsiveness to operator commands.

Acknowledgments. We would like to thank to Program Pendanaan Inovasi 2024 Skema Program Pengembangan Produk (P4) Universitas Indonesia with PRIMESTeP Grant from Ministry of Education, Culture, Research and Technology, Republic of Indonesia for partially supported this research.

References

1. Radzi MF, Saputra HM, Baskoro CHAHB, Mirdanies M, Aziz RIA, Ismail N (2021) PID-based fraction of inspired oxygen (FiO_2) control system in ICU ventilator. In: 2021 IEEE 7th

international conference on smart instrumentation, measurement and applications ICSIMA 2021, pp 185–189. https://doi.org/10.1109/ICSIMA50015.2021.9525942
2. Chikata Y, Ohnishi S, Nishimura M (2017) Humidity and inspired oxygen concentration during high-flow nasal cannula therapy in neonatal and infant lung models. Respir Care 62:532–537. https://doi.org/10.4187/respcare.05319
3. Chandel A et al (2020) High-flow nasal cannula in COVID-19: outcomes of application and examination of the ROX index to predict success. Respir Care 66:909–919. https://doi.org/10.4187/RESPCARE.08631
4. Samolski D, Antón A, Güell R, Sanz F, Giner J, Casan P (2006) Inspired oxygen fraction achieved with a portable ventilator: determinant factors. Respir Med 100:1608–1613. https://doi.org/10.1016/J.RMED.2005.12.005
5. Coudroy R, Frat JP, Girault C, Thille AW (2020) Reliability of methods to estimate the fraction of inspired oxygen in patients with acute respiratory failure breathing through non-rebreather reservoir bag oxygen mask. Thorax 75:805–807. https://doi.org/10.1136/THORAXJNL-2020-214863
6. Claure N, Bancalari E (2013) Automated closed loop control of inspired oxygen concentration. Respir Care 58:151–161. https://doi.org/10.4187/RESPCARE.01955
7. Dani C (2019) Automated control of inspired oxygen (FiO_2) in preterm infants: literature review. Pediatr Pulmonol 54:358–363. https://doi.org/10.1002/PPUL.24238
8. Chanyagorn P, Kiratiwudhikul P (2016) Automatic control of fraction of inspired oxygen in neonatal oxygen therapy using fuzzy logic control. IEIE Trans Smart Process Comput 5:107–116. https://doi.org/10.5573/IEIESPC.2016.5.2.107
9. Fajrin HR, Afifah SK, Wibowo SA (2021) Mixed pressure control innovation of oxygen and air pressure in ventilator with safety valve. J Phys Conf Ser 1933:012106. https://doi.org/10.1088/1742-6596/1933/1/012106
10. Alfaray RI, Mahfud MI, Faizun RS (2019) Duration of ventilation support usage and development of ventilator-associated pneumonia: when is the Most time at risk? Indones J Anesth Reanim 1:26–31. https://doi.org/10.20473/IJAR.V1I12019.26-31
11. Saseedharan S (2018) Modes of mechanical ventilation. MCQs Mech Vent Case-Based Approach 25–25. https://doi.org/10.5005/JP/BOOKS/14169_3
12. Fan C, Zheng J, Yang Y, Si B (2017) Design of respiratory parameter estimator based on linear RCC model. In: 2017 4th International conference on systems and informatics, ICSAI 2017. Jan 2018, pp 94–98. https://doi.org/10.1109/ICSAI.2017.8248270
13. Morton SE et al (2020) Virtual patient Modeling and prediction validation for pressure controlled mechanical ventilation. IFAC-PapersOnLine 53:16221–16226. https://doi.org/10.1016/J.IFACOL.2020.12.615
14. Pratama SA, Saputra HM, Hamidi EAZ, Baskoro CHAHB, Haniff M, Mirdanies M (2022) PI control of pressure inspiration in ICU ventilator. In: Proceeding 2022 8th international conference on wireless and telematics ICWT 2022. https://doi.org/10.1109/ICWT55831.2022.9935456
15. Amaranggana A, Priramadhi R (2021) Air pressure control on ventilator using fuzzy logic method. In: Proceedings—2021 4th international conference of computer and informatics engineering. IT-based digital industrial innovation for the welfare society, IC2IE 2021, pp 158–163. https://doi.org/10.1109/IC2IE53219.2021.9649324
16. Morales S et al (2021) Pressure and volume control of a non-invasive mechanical ventilator: a PI and LQR approach. In: 2021 9th International conference on control, mechatronics and automation ICCMA 2021, pp 67–71. https://doi.org/10.1109/ICCMA54375.2021.9646186
17. Yamaguchi K, Toda H (2022) Development of a slider type mechanical ventilator system using ambu bag for emergency breathing assistance. In: 2022 6th International conference on automation, control and robots ICACR 2022. pp 88–92. https://doi.org/10.1109/ICACR55854.2022.9935535

18. Aziz A et al (2023) Revolutionizing critical care: a smart positive pressure based non-invasive ventilator with real-time iot health monitoring system. In: Proceeding 2023 IEEE 6th international conference on knowledge innovation and invention ICKII 2023. pp 1–6. https://doi.org/10.1109/ICKII58656.2023.10332616
19. Arshad M, Mehmood K, Lazoglu I (2023) Development of a non-invasive ventilator for emergency and beyond. Comput Biol Med 167:107670. https://doi.org/10.1016/J.COMPBIOMED.2023.107670
20. Kumar M et al (2024) Design and technical evaluation of an AMBU-BAG based low-cost ventilator-AARMED. Smart Heal 31:100445. https://doi.org/10.1016/J.SMHL.2023.100445
21. Pristianto EJ et al (2021) High-flow and high-pressure oxygen mixing for ventilator system. In: Proceeding—2021 International Conference Radar, Antenna, Microwave, Electron. Telecommun. Manag. Impact Covid-19 Pandemic Together Facing Challenges Through Electron. ICTs, ICRAMET 2021. pp 201–204. https://doi.org/10.1109/ICRAMET53537.2021.9650464
22. Basari, Abuzairi T, Tjahjadi H, Gunawan I (2024) Ventilator yang Mudah Dipindahkan dan Digunakan dalam Keadaan Darurat Berbasis Sistem Pneumatik, Patent Granted No: IDP000096777, Ministry of Law and Human Rights, Republic of Indonesia. https://pdki-indonesia.dgip.go.id/. Last accessed 04 Dec 2024
23. VT650 gas flow analyzer|Fluke biomedical (2022). https://www.flukebiomedical.com/products/biomedical-test-equipment/gas-flow-analyzers-ventilator-testers/vt650-gas-flow-analyzer-ventilator-tester. Last accessed 05 July 2022
24. STM32F411CEU6—WeAct black pill V2.0|STM32-base project (2024). https://stm32-base.org/boards/STM32F411CEU6-WeAct-Black-Pill-V2.0.html. Last accessed 02 Oct 2024
25. STM32F407VGT6—STM32F4XX M|STM32-base project (2024). https://stm32-base.org/boards/STM32F407VGT6-STM32F4XX-M.html. Last accessed 02 Oct 2024
26. Indonesia DJPK (2018) Metode Kerja Pengujian dan Kalibrasi Alat Kesehatan:Decree No. HK.02.02/V/5771/2018 28 Nov. 2018 Ministry of Health, Republic of Indonesia, ISBN 9786024165000. https://repository.kemkes.go.id/book/437. Last accessed 02 Oct 2024
27. Bishop M (2022) Basic principles of mechanical ventilation, Chapter 5 - Control modes: initial ventilator settings, Sault College, 2022. https://melodybishoprt.com/resources/. Last accessed 02 Oct 2024
28. Kallet RH, Campbell AR, Alonso JA, Morabito DJ, Mackersie RC (2000) The effects of pressure control versus volume control assisted ventilation on patient work of breathing in acute lung injury and acute respiratory distress syndrome. Respir Care 45(11). https://pubmed.ncbi.nlm.nih.gov/10980100/. Last accessed 02 Oct 2024
29. Lain DC, DiBenedetto R, Morris SL, Nguyen A, Saulters R, Causey D (1989) Pressure control inverse ratio ventilation as a method to reduce peak inspiratory pressure and provide adequate ventilation and oxygenation. Chest 95:1081–1088. https://doi.org/10.1378/CHEST.95.5.1081
30. Mlcak RP, Suman OE, Sousse LE, Herndon DN (2018) Respiratory care. Total burn care, fifth ed. pp 195–204.e2. https://doi.org/10.1016/B978-0-323-47661-4.00018-6
31. Bishop M (2022) Basic principles of mechanical ventilation, Chapter 2 - Oxygenation and the fundamentals of mechanical ventilation, Sault College, 2022. https://melodybishoprt.com/resources/. Last accessed 02 Oct 2024
32. Fuentes S, Chowdhury YS (2022) Fraction of inspired oxygen. StatPearls. https://www.ncbi.nlm.nih.gov/books/NBK560867/. Last accessed 02 Oct 2024
33. Mora Carpio AL, Mora JI (2022) Assist-control ventilation. StatPearls. https://www.ncbi.nlm.nih.gov/books/NBK441856/. Last accessed 02 Oct 2024
34. Allardet-Servent J, Sicard G, Metz V, Chiche L (2019) Benefits and risks of oxygen therapy during acute medical illness: just a matter of dose! La Rev Med interne 40:670–676. https://doi.org/10.1016/J.REVMED.2019.04.003

35. Chu DK et al (2018) Mortality and morbidity in acutely ill adults treated with liberal versus conservative oxygen therapy (IOTA): a systematic review and meta-analysis. Lancet (Lond, Engl) 391:1693–1705. https://doi.org/10.1016/S0140-6736(18)30479-3
36. Borrello M (2005) Modeling and control of systems for critical care ventilation. Proc Am Control Conf 3:2166–2180. https://doi.org/10.1109/ACC.2005.1470291
37. Paru-Paru UjilDraeger (2024). https://www.draeger.com/id_id/Products/Draeger-Self-Test-Lung. Last accessed 03 Nov 2024
38. Aziz SA, Basari (2024) Uji Kinerja Ventilator pada Mode Synchronized Intermittent Mandatory Ventilation (SIMV). Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika (ELKOMIKA) 12(3). https://doi.org/10.26760/elkomika.v12i3.656
39. How to calculate percent error (2024). https://www.thoughtco.com/how-to-calculate-percent-error-609584. Last accessed 25 Oct 2024
40. Dräger. Ventilator Dräger Savina® 300. https://www.draeger.com/id_id/Products/Savina-300. Last accessed 02 Oct 2024
41. GE Healthcare. Ventilator Carescape R860 Specification. https://www.gehealthcare.com/. Last accessed 02 Oct 2024
42. Hamilton-C1 ICU mechanical ventilator (2024). https://www.hamilton-medical.com/en/Products/HAMILTON-C1.html. Last accessed 02 Oct 2024

Evaluating Sand Resources for Radiation Protection: A Study of Silica and Iron Sand Distribution Along Yogyakarta Beach

Muhammad Fakhrurreza[1,2(✉)], Anisa Nur Istiqomah[1], and Rajendra Andita Nugraha[1]

[1] Radiology Program, Faculty of Health Sciences, Universitas Aisyiyah Yogyakarta, Jl. Siliwangi No 63, DI Yogyakarta 55292, Indonesia
muhammadfakhrurreza@unisayogya.ac.id

[2] Physic Department, Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, DI Yogyakarta 55281, Indonesia

Abstract. Radiation shielding, time- and distance minimization, and minimization of radiation are the three methods used in external radiation protection to regulate radiation acceptance. Boundary walls are materials other than lead, steel, and other heavy materials that are used to make external radiation shields. Boundary walls are frequently constructed using bricks, particularly in structures belonging to the Radiology Installation. This study set out to evaluate the sand quality of Yogyakarta Beach, specifically identifying the areas with the highest concentrations of silica and iron sand. Direct sampling along the path around Yogyakarta Beach. taken at 11 places, with a 3–4 km distance between each location point and the beginning point for the sand sampling at Parangendog beach, and extending up to 35 km to Congot beach. Five kilogram samples of each spot were taken, and a geomagnetic device was used to separate the silica and iron sand. Following the separating procedure, the mass of each sand was weighed. An average was then calculated per five kilos of both silica and iron sand, and a graphic was created to compare the two. Statistical analysis using Wilcoxon Sign Rank Test. According to the study's results, kilometers 10–25 had the highest concentration of iron sand. At Cubung Jangkang Beach, kilometer 25 has the highest average iron sand content (0.77 kg), followed by kilometer 21 at Bugel Beach (0.757 kg) and kilometer 14 at Kuwaru Beach (0.751 kg). Thus, Kuwaru Beach, Bugel Beach, and Cubung Jangkang Beach are appropriate to use as raw materials for anti-radiation bricks. The p value <0.05 so that there is a significant difference in the average content of iron sand and silica. The study is expected to be the basis for making bricks from iron sand with the material used coming from local sand from the Yogyakarta beach.

Keywords: Iron sand · Brick Anti-radiation · Radiation Shielding

1 Introduction

The use of x-rays in nuclear medicine is one way to improve public health. This application has been quite diverse ranging from radiation for diagnostics, dental x-ray examinations and x-ray radiation for therapy [1, 2]. An estimated 2500 million diagnostic

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 446–458, 2026.
https://doi.org/10.1007/978-981-95-9233-3_32

radiological exams, 32 million nuclear medicine exams, and 5.5 million radiotherapy sessions are conducted globally year [3]. Radiation technology improves life, but depending on the length, degree, and location of exposure, it can also cause a variety of health issues [4].

Effective shielding materials are needed between the radiation source and radiation officers in order to protect the environment and radiation officers from the harmful effects of exposure to leaked, scattered, or direct ionizing radiation, such as γ rays and X-rays [5–7]. Based on Nuclear Energy Regulatory Agency of Indonesia No 4 year 2020 planning for radiation protection needs to be implemented at the stages of facility planning, design, and use for radiology installations as well as for any building that uses radiation power for operational purposes.

The most common element used for protection off radiation is lead. However, there are certain limitations with Pb, like its toxicity and high cost. Thus, it is necessary to locate alternative materials that may be employed without limitations for protecting toward field radiation [1, 8]. According to tests, materials like granite [9] and ceramics (bricks) [10] have a linear attenuation coefficient μ of 0.08 to 0.13 cm^{-1} for a 60 Co gamma ray source, which is even lower than the 0.15–0.19 of various heavy concrete combinations.

An iron-rich material was used to create a reasonably priced radiation shielding brick with a comparatively high density. When compared to regular concrete, clay and iron minerals have superior radiation properties. Bricks are made from a variety of materials and are regarded as one of the fundamental parts of a building [11, 12]. Mori et. al [13] uses Iron (III) Oxide-Based Ceramic Material as an anti-radiation material. For high-energy gamma rays, the ceramic's measured attenuation coefficient ranged from 0.268 to 0.355, or around 40% less than that of lead and twice that of concrete. The combination of iron powder mixture with mortar provides the best protection against gamma rays and neutrons [14, 15]. For a material thickness of 2 cm and for low X-ray radiation energy, radiation shielding material constructed from a combination of fly ash and Fe_2O_3 had good X-ray radiation energy absorption [16]. We can conclude from the previously described research that iron is a good substitute for other materials to produce anti-radiation bricks.

In Indonesia, more precisely in the province of Yogyakarta, there is a beach on the south side that stretches along Yogyakarta and faces directly onto the Indian Ocean. The beach sand in Yogyakarta contains iron sand that can be used as a basic material for anti-radiation bricks. According to Fakhrurreza [17], Selatan beach sand have HVL 0.8644 cm and Trisik beach sand with a value of 1.314 cm^{-1} [18]. However, the quantity of silica and iron sand along the Yogyakarta coast has not been studied. This study aimed to determine the average concentration of iron and silica sand throughout a 35-km stretch of the Yogyakarta coast. Samples were collected at 11 locations spaced 3–4 km apart.

The objective of this investigation is to determine the location of the highest concentration of iron sand along the coast of Yogyakarta by calculating the average value of iron sand and silica content at each location. The data analysis of this investigation is restricted to the use of a homemade geomagnetic tool to separate iron sand and silica and to determine the quantity of iron sand. Certainly, this presents a challenge, as no molecular analysis was conducted. This study is anticipated to serve as the foundation

for the production of anti-radiation bricks by utilizing the beach point with the maximum iron sand content along the coast of Yogyakarta.

2 Radiation Shielding

Shielding is a critical material that is employed to prevent radiation from reaching the general public and radiation officers. Proper shielding between the radiation source and the individual further reduces the external exposure dose to acceptable levels. A temporary shielding such as shielding assemblies, lead bricks, paraffin blocks, portable shields, etc. required for specific tasks. A proper thickness of shielding is needed by considering the type of radiation (X-rays, gamma, neutrons, etc.) and shielding material to be used (lead, concrete, steel, paraffin, etc.) and the greater energy the thicker shield.

Substances characterized by high atomic numbers and density demonstrate superior photon absorption efficiency. The required thickness of shielding is typically established through the concept of half-value layer (HVL) or half-value thickness (HVT) [19]. HVL denotes the specific material thickness that decreases the air kerma of X-ray or gamma radiation to 50% of its original intensity [20]. This value, which is measured in millimeters or centimeters, is dependent on photon energy and inversely correlated with the attenuation coefficient. It is calculated by applying and solving the attenuation equation for x, as expressed in Eq. (1):

$$\frac{1}{2} = e^{-\mu x_{1/2}}$$

$$x_{\frac{1}{2}} = \frac{\ln(2)}{\mu}$$

$$x_{\frac{1}{2}} = \frac{0.693}{\mu} = HVL \tag{1}$$

Shielding calculations for X-rays may utilize the inherent technical specifications provided by the manufacturer, such as tube type, generator voltage, and the filtration material and thickness. The barrier transmission factor, B—defined as the fraction of radiation that penetrates a protective barrier at a given distance, d, from the source due to primary, scattered, or leakage radiation—can be determined using Eq. (2):

$$B = Pd^2/WUT, \tag{2}$$

In this context, P represents the maximum allowable dose rate (air kerma) for both controlled and uncontrolled areas. For radiation professionals such as radiographers and radiologists, this limit is generally 1 mSv/week, whereas for members of the public it is 0.02 mSv/week. W refers to the workload in Gy per week, U to the use factor, and T to the occupancy factor, which quantifies the proportion of time a location is occupied while the X-ray beam is active. These values enable estimation of the unshielded kerma and calculation of the transmission factor required to achieve P/T (Eq. (3)):

$$K_0 = \frac{K^1 UN}{d^2}, \tag{3}$$

In Eq. (3), K^1 denotes the average kerma (both primary and secondary) expected during a patient procedure, U is the use factor (set to 1 for secondary barriers), N is the number of procedures performed weekly, and d is the distance from the source. The appropriate thickness, x, of a protective barrier is therefore that which results in a transmission factor, B, not exceeding the permissible value (Eq. 4):

$$B_{(x)} = \frac{P}{T}/K_0 \tag{4}$$

Furthermore, practical shielding models should account for additional attenuation already provided by existing structures, such as the image receptor (cassette or holder) and the examination table (Eq. 5),

$$x_{tot} = x + x_{pre} \tag{5}$$

3 Materials and Methods

3.1 Raw Sand Materials

Sand samples were collected at each of the 11 pre-selected beach spots, which are spaced out over a distance of 3–4 km to the west. To make it easier for researchers to locate beach sites, points were identified by using the Google Maps application to determine the position for sand sample collection. The first point is 35 km away, starting from Parangendog Beach with coordinate −8.0278372, 110.3398002 and ending at Congot Beach. Table 1 displays the locations of the beaches. Five kilogram of sand was randomly selected from the shoreline area at each beach point and placed in sacks.

Table 1. Beach sand sampling point

No	Beach name	km	Amount of sample
1	Parangendog	0	5 kg
2	Cemara Sewu	4	5 kg
3	Depok	7	5 kg
4	Samas	10	5 kg
5	Kuwaru	14	5 kg
6	Trisik	18	5 kg
7	Bugel	22	5 kg
8	Cubung Jangkang	25	5 kg
9	Palma	28	5 kg
10	Karang Wuni	31	5 kg
11	Congot	35	5 kg

3.2 Raw Sand Processing

In order to get rid of the salt in the beach sand, the raw sand is first cleaned. After that, the sand is sun-dried until it is completely dry, then it is filtered through a coconut milk filter to remove any remaining dirt or pebbles. After being cleaned, the 5 kg of sand from one area are separated into 5 samples, each weighing 1 kg. As shown in Fig. 1, 1 kg of sand was separated into 10 raw sand samples, each weighing 100 g. Every collection of sand is gathered in a container based on every location point.

(a) (b)

Fig. 1. Raw sand processing (a) raw sand samples; (b) grouping of Yogyakarta beach sand samples

3.3 The Process of Separating Silica and Iron Sand

The geomagnetic method, which divides rocks with magnetic properties from those without, is used to separate silica sand and iron sand. Using tool similar to those in Fig. 2, researchers independently created a geomagnetic tool for this investigation.

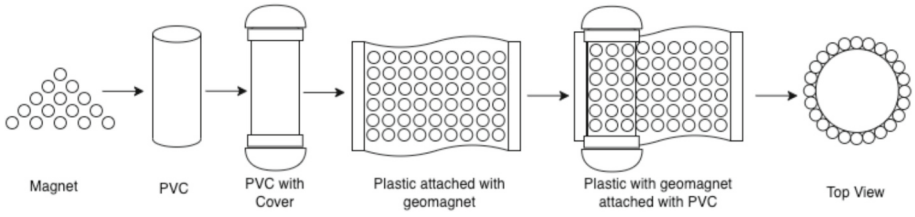


Fig. 2. Geomagnetic method tool

The tool is made using 2½ inch PVC pipe with a length of 30 cm. The mica book is cut and glued using a glue gun, on the right and left edges of the mica book, adhesive is attached with a glue gun and sewn manually. Then the magnet is installed on the outside of the pipe or attached to the mica book with a distance of 2 cm each. The installation of the magnet is attached using a glue gun. The outer magnet is coated with tape so that the magnet does not come off easily. One end of the pipe is closed so that sand can be

inserted into the pipe. At the end of the pipe there is a cover as a place for the separated material to exit. The results of making the geomagnetic tool are shown in Fig. 3.

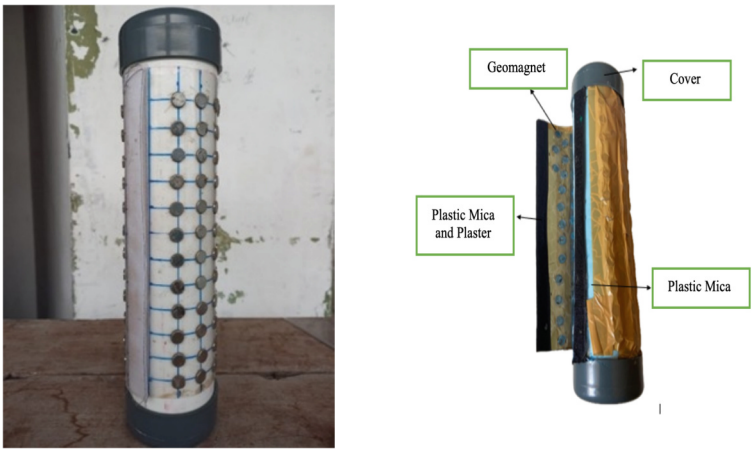


Fig. 3. Geomagnetic separator tool

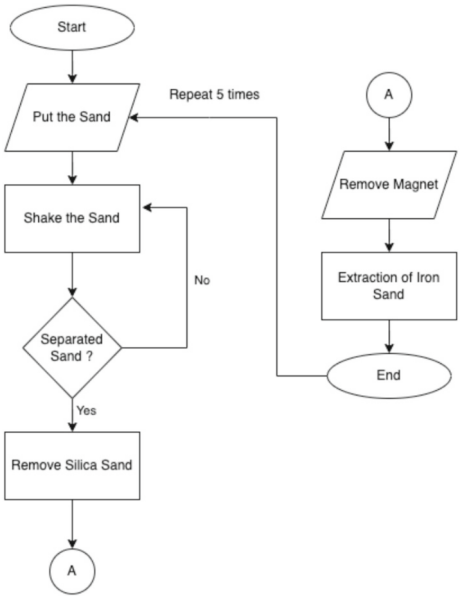


Fig. 4. Flowchart of separation of iron sand and silica sand using geomagnetic tool

To separate the two types of sand by preparing 2 containers for both types of sand. Flowchart shown in Fig. 4. A 100 g sand sample is inserted into the PVC pipe and both

ends of the pipe are closed. The pipe is rotated for 1 min. Then open the top cover and the sand is removed for the sand that does not contain iron into the prepared container. The magnet surrounding the pipe is released so that the iron sand can be released. The sand containing iron is poured into another container. The process is repeated 5 times each for both iron sand and silica sand, so that the iron sand obtained is pure iron sand.

3.4 Data Analysis

Sand after going through the separation process per 100 g obtained the results of iron sand and silica, then weighed each iron sand and non-iron then the results of the data obtained will appear. Furthermore, the data of 100 g of sand becomes 5 kg and each iron and non-iron sand is taken on average. From the results of the data, the data is processed using statistics to determine the difference in the average content of iron sand and silica sand using Python with Google Colab Notebook with the Wilcoxon Sign Rank Test. A diagram is made and a comparison is made between iron and non-iron sand of Yogyakarta Beach. Finally, observations are made using a digital microscope, both in terms of shape, size and color.

4 Results and Discussion

4.1 Average Measurement of Iron and Silica Sand

From all the data obtained, it can be seen that the average weight of iron sand and silica sand from Yogyakarta Beach from kilometer 0 to kilometer 35 is presented in the Fig. 5. The data is the result of the average sand per 5 kg of samples at each beach point. The average is taken from the number of 50 sand samples separated using a geomagnetic tool.

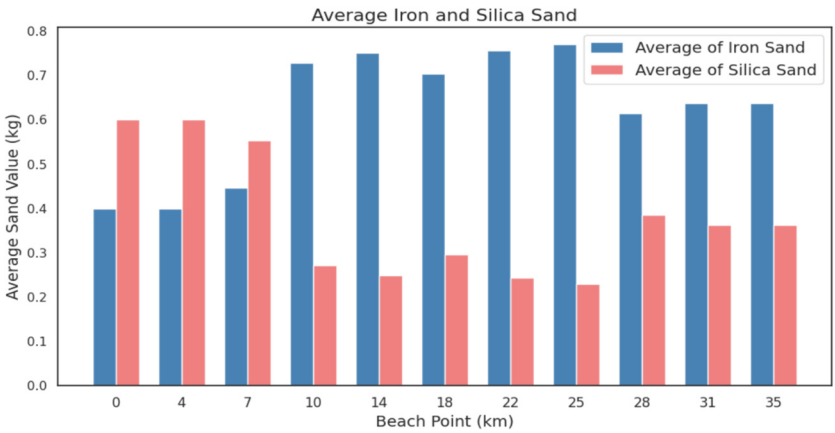


Fig. 5. Average measurement of iron and silica sand

The value of iron sand and silica sand shows fluctuations or ups and downs. The Iron Sand content from kilometer 0 to kilometer 7 has almost the same value, while at kilometer 10 to kilometer 25 it increases with the highest value at kilometer 25 of 0.77 kg. However, after further analysis at kilometer 28 to kilometer 35 there is a decrease in value. Meanwhile, for silica sand from kilometer 0 to kilometer 35 it also increases and decreases. The highest value at kilometer 0 is 0.4 kg. At kilometer 10 to kilometer 25 it has almost the same value, but after further analysis at kilometer 7 to kilometer 25 it decreases with the lowest value at kilometer 25 and at kilometers 28–35 it has almost the same value.

Figure 5 illustrates how this study establishes that the average iron and silica sand composition varies from kilometer 0 to kilometer 35. At an average of 0.7704 kg kilometer 25 Cubung Jangkang beach, kilometer 22 Bugel beach, and kilometer 14 Kuwaru beach have the greatest iron sand contents, respectively, at 0.757 and 0.751 kg. Furthermore, kilometer 4 Cemoro Sewu beach has the lowest average iron sand percentage, at 0.3992.

Table 2. Wilcoxon sign rank test

Test	Statistic	ρ-value
Wilcoxon Sign Rank Test	6.0	0.013671875

The author subsequently performed a Wilcoxon signed-rank test to evaluate the average content of iron sand and silica sand, the result are as in the Table 2. The results indicated a p value of <0.05, signifying a significant difference in the average concentrations of iron sand and silica at 11 coastal locations.

Beach sand can be utilized with extra caution when used as raw material for bricks; first, wash the sand to remove any salt residue [18]. Iron (III) oxide-based ceramic brick has a lower shielding effect than lead, but it is still non-toxic, rust-proof, and sufficiently durable even at high temperatures. It is also simple to produce into a variety of shapes [13].

The iron sand content in the three beaches above is quite a lot so that it is very possible to be used as a basic material for anti-radiation bricks. However, the separation of iron sand and silica still uses manual methods so that to get more accurate values, better geomagnetic tools can be used.

4.2 Characteristics of Iron Sand and Silica Sand

The sand on the Yogyakarta beach at kilometer 0–7 is different from the sand at kilometer 7 and above because the sand on the beach contains white gravel grains that come from the erosion of rocks in the beach area. Then the iron sand content of the beach at kilometer 0–7 is lower than the beach at kilometer 7 and above. Figure 6 shows an image of iron sand and silica viewed using a microscope with a magnification of 1600x. On the sand of Yogyakarta beach kilometer 10–25 is different from the sand of kilometer 0–7 and kilometer 25–35 where on the beach the sand is predominantly containing black gravel

grains originating from a mixture of river water from the Opak river and the Progo river. Hence, the iron sand content of the beach kilometer 10–25 is higher than other beaches. At kilometer 28–35 is different from the sand below kilometer 28. Because the sand on this beach is mixed with pine tree branches. Hence, the iron sand content on the beach at kilometer 28–35 is lower than the sand on the beach at kilometer 10–25.

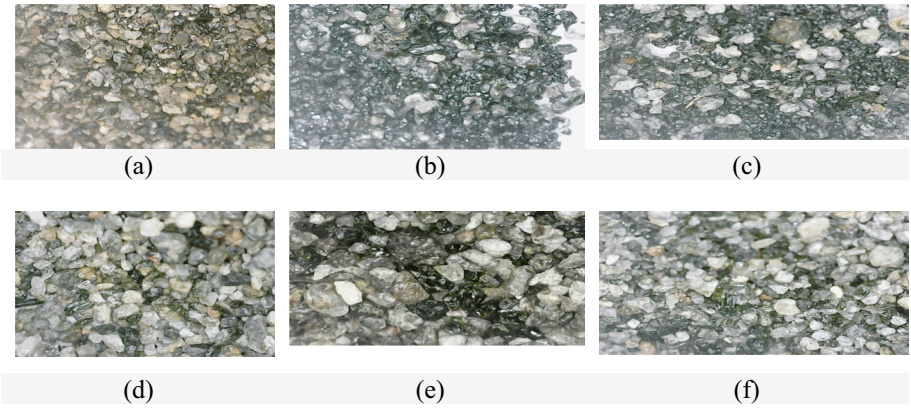


Fig. 6. (a) Iron sand KM 0-7; (b) iron sand KM 10-25; (c) iron sand KM 28-35; (d) iron silica KM 0-7; (e) iron silica KM 10-25; (f) iron silica KM 28-35

When using ionizing radiation in spatial planning, walls must be thick enough to meet Nuclear Energy Regulatory Agency (BAPETEN) standards and be covered in materials with high atomic numbers to ensure optimal absorption of scattered radiation [21].

Iron is one among them; its density of 55.845 μ makes it capable of efficiently absorbing dispersed radiation. The material's capacity to block and absorb radiation energy increases with density. Iron was chosen as the primary material for this investigation due to its accessibility and low cost in comparison to other materials with comparable qualities. In addition to being simple to work with and shape, iron has the ability to improve gamma ray attenuation [22].

The results of observations using a digital microscope show that beaches with a higher iron sand content tend to be blacker than beaches with a lower iron content (Table 3). In general, we can see the difference between iron sand and silica sand but have not yet done physical characterization using more sophisticated tools such as SEM. However, it is enough to see as an initial study of making anti-radiation bricks.

Table 3. Comparison of radiation shielding materials

Material type	Concrete/mortar	Findings	Ref.
Magnetite	Ultra-high performance concrete	The use of magnetite rocks as a fine aggregate has proven to enhance the radiation shielding performance of UHPHSC more effectively.	[23]
Siderurgical aggregate	Concrete	The incorporation of siderurgical aggregates into concrete formulations enhances the effective atomic number and improves the mass attenuation coefficient.	[24]
Steel furnace slag aggregate	High Strength Concrete	The volume percentage of coarse aggregate and the binder-to-fine-to-coarse aggregate ratio produced a dense and well-compacted concrete matrix, enhancing its density and improving its capacity to attenuate γ-rays, so rendering it an efficient radiation shield	[25]
Nano Ferrosilicon and magnetite	Ultra-high performance heavyweight concrete	The combination of 100% magnetite and 3% nano ferrosilicon demonstrated its efficacy in radiation shielding, as evidenced by measurable outcomes, when exposed to a radiation source with an energy level of 1250 KeV.	[26]
WO3 and barite	Concrete	The experimental results showed significant alterations in the concrete samples' linear attenuation coefficient (LAC)	[27]
Recycled coarse aggregate and steel slag	Concrete	Regardless of the kind of aggregate or the cement paste's chemical makeup, attenuation is enhanced to a larger extent with increasing unit weight.	[28]

5 Discussion

The results showed significant variations in iron sand and silica content along the Yogyakarta Coast. The highest concentration of iron sand was found in the area between kilometers 10 to 25, indicating the great potential of this area as a source of raw materials for making anti-radiation bricks. Statistical analysis confirmed significant differences in the distribution of these minerals. However, further characterization using SEM is needed to understand the physical properties of iron sand in depth. The potential use of iron sand as a raw material for anti-radiation bricks needs to be further developed, especially the value of its linear attenuation coefficient or HVL to determine protection toward radiation.

There are several studies related to the materials for making radiation shielding materials as in Table 3. Materials such as Magnetite [23], Siderurgical aggregate [24], Steel furnace slag aggregate [25], Nano Ferrosilicon [26], WO_3 [27], Barite [27], recycled coarse aggregate and steel slag [28] can be used as basic materials for concrete for radiation shielding. However, iron sand can be considered as a basic material for radiation shielding considering its ability to withstand radiation and have high density [14, 15].

This study has limitations, in the sample processing still using a simple method, using handmade geomagnet tool and limited to the amount of iron sand content at certain beach points. This study, it is expected to be developed for making bricks or concrete with iron sand as the basic material taken locally on the Yogyakarta coast—Cubung Jangkang beach, Bugel beach and Kuwaru beach. This research is beneficial for the replacement of lead-based wall coatings with anti-radiation walls that do not require lead coating.

6 Conclusion

The aim of this study is to identify the location with the highest concentration of iron sand along the Yogyakarta coast. Beach sand is one of the good basic materials for radiological buildings because it contains high iron sand so it can withstand radiation exposure. The best iron sand content for the basic material for making anti-radiation brick buildings is taken at kilometers 10–25 in the Kuwaru beach area to Cubung Jangkang beach which can be used for further research. Three beaches—Cubung Jangkang Beach, Bugel Beach, and Kuwaru Beach—have the highest concentration of iron sand with the value 0.7704 kg, 0.757 kg and 0.751 kg respectively. There are significant difference in the average concentrations of iron sand and silica at 11 coastal locations with p-value 0.01367 from Wilcoxon Sign Rank Test. Future studies will focus on using iron sand from three coastal locations with the highest iron sand content as the basic material for bricks.

References

1. Issa SAM, Tekin HO, Elsaman R, Kilicoglu O, Saddeek YB, Sayyed MI (2019) Radiation shielding and mechanical properties of Al_2O_3-Na_2O-B_2O_3-Bi_2O_3 glasses using MCNPX Monte Carlo code. Mater Chem Phys 223:209–219. https://doi.org/10.1016/j.matchemphys.2018.10.064

2. Lee C-M, Lee YH, Lee KJ (2007) Cracking effect on gamma-ray shielding performance in concrete structure. Prog Nucl Energy 49:303–312. https://doi.org/10.1016/j.pnucene.2007.01.006
3. Lakhwani OP, Dalal V, Jindal M, Nagala A (2019) Radiation protection and standardization. J Clin Orthop Trauma 10:738–743. https://doi.org/10.1016/j.jcot.2018.08.010
4. Johnson TE (2017) Introduction to health physics. McGraw Hill Professional
5. Lakshminarayana G et al (2018) Vibrational, thermal features, and photon attenuation coefficients evaluation for TeO_2-B_2O_3-BaO-ZnO-Na_2O-Er_2O_3-Pr_6O_{11} glasses as gamma-rays shielding materials. J Non-Cryst Solids 481:568–578. https://doi.org/10.1016/j.jnoncrysol.2017.11.049
6. Singh VP, Badiger NM, Chanthima N, Kaewkhao J (2014) Evaluation of gamma-ray exposure buildup factors and neutron shielding for bismuth borosilicate glasses. Radiat Phys Chem 98:14–21. https://doi.org/10.1016/j.radphyschem.2013.12.029
7. Akman F, Durak R, Turhan MF, Kaçal MR (2015) Studies on effective atomic numbers, electron densities from mass attenuation coefficients near the K edge in some samarium compounds. Appl Radiat Isot 101:107–113. https://doi.org/10.1016/j.apradiso.2015.04.001
8. Issa SAM (2016) Effective atomic number and mass attenuation coefficient of PbO–BaO–B_2O_3 glass system. Radiat Phys Chem 120:33–37. https://doi.org/10.1016/j.radphyschem.2015.11.025
9. Najam LA, Hashim AK, Ahmed HA, Hassan IM (2016) Study the attenuation coefficient of granite to use it as shields against gamma ray. Detection 04:33–39. https://doi.org/10.4236/detection.2016.42005
10. Jawad AA, Demirkol N, Gunoğlu K, Akkurt I (2019) Radiation shielding properties of some ceramic wasted samples. Int J Environ Sci Technol 16:5039–5042. https://doi.org/10.1007/s13762-019-02240-7
11. Mann HS, Brar GS, Mudahar GS (2016) Gamma-ray shielding effectiveness of novel light-weight clay-flyash bricks. Radiat Phys Chem 127:97–101. https://doi.org/10.1016/j.radphyschem.2016.06.013
12. Mahmoud KA, Tashlykov OL, Mhareb MHA, Almuqrin AH, Alajerami YSM, Sayyed MI (2021) A new heavy-mineral doped clay brick for gamma-ray protection purposes. Appl Radiat Isot 173:109720. https://doi.org/10.1016/j.apradiso.2021.109720
13. Mori H, Oku Y, Mannami Y, Kunisada T (2020) Iron (III) oxide-based ceramic material for radiation shielding. Ceramics 3:258–264. https://doi.org/10.3390/ceramics3020023
14. Al-Saleh WM, Elsafi M, Almutairi HM, Nabil IM, El-Nahal MA (2024) A comprehensive study of the shielding ability from ionizing radiation of different mortars using iron filings and bismuth oxide. Sci Rep 14. https://doi.org/10.1038/s41598-024-60188-2
15. Dahlan K, Haryati E, Aninam YS (2018) The use of Papuan iron sand and river sand for fine aggregate in mortar for nuclear radiation shield application. J Phys Conf Ser 997. https://doi.org/10.1088/1742-6596/997/1/012023
16. Budiawan MA, Suryani S, Abdullah B, Tahir D (2019) Analysis of absorption properties of a composite FlyAsh and Fe_2O_3 for X-ray radiation shielding applications. IOP Conf Ser Mater Sci Eng 593:3–7. https://doi.org/10.1088/1757-899X/593/1/012014
17. Fakhrurreza M, Astari FM (2019) Anti radiation building design: local sand type analysis to get concrete X-radiation resistant X-ray. Jurnal Imejing Diagnostik (JImeD) 5:25. https://doi.org/10.31983/jimed.v5i1.4001
18. Fakhrurreza M, Astari FM, Susilo SH (2023) Anti-radiation building design: consistency of linear attenuation coefficient in black sand South Beach Diy Province to get the best X-ray radiation resistance concrete. Atlantis Press International BV. https://doi.org/10.2991/978-94-6463-190-6_10
19. Domenech H (2016) Shielding. In: Radiation safety management and programs. Springer Cham, Switzerland, pp 97–109. https://doi.org/10.1007/978-3-319-42671-6_7

20. Tekin HO, ALMisned G, Issa SAM, Zakaly HMH (2022) A rapid and direct method for half value layer calculations for nuclear safety studies using MCNPX Monte Carlo code. Nuclear Eng Technol 54:3317–3323. https://doi.org/10.1016/J.NET.2022.03.037
21. Dasril DN, Dewilza N (2020) Uji Efektifitas Dinding Ruangan Panoramik Instalasi Radiologi RSUD Prof. Dr. MA Hanafiah SM Batusangkar Menggunakan TLD-100. Phys Educ Res J 2:95. https://doi.org/10.21580/perj.2020.2.2.5087
22. Alkarrani H, ALMisned G, Tekin HO (2024) Assessing the efficacy of some heavy-metal infused concrete mixtures in gamma-ray and neutron shielding applications. Radiat Phys Chem 223:111988. https://doi.org/10.1016/j.radphyschem.2024.111988
23. Zeyad AM, Hakeem IY, Amin M, Tayeh BA, Agwa IS (2022) Effect of aggregate and fibre types on ultra-high-performance concrete designed for radiation shielding. J Build Eng 58:104960. https://doi.org/10.1016/J.JOBE.2022.104960
24. Tamayo P, Thomas C, Rico J, Pérez S, Mañanes A (2022) Radiation shielding properties of siderurgical aggregate concrete. Constr Build Mater 319:126098. https://doi.org/10.1016/J.CONBUILDMAT.2021.126098
25. Khalaf MA, Ban CC, Ramli M, Ahmed NM, Sern LJ, Khaleel HA (2020) Physicomechanical and gamma-ray shielding properties of high-strength heavyweight concrete containing steel furnace slag aggregate. J Build Eng 30:101306. https://doi.org/10.1016/J.JOBE.2020.101306
26. Heniegal AM, Amin M, Nagib SH, Youssef H, Agwa IS (2022) Effect of nano ferrosilicon and heavyweight fine aggregates on the properties and radiation shielding of ultra-high performance heavyweight concrete. Case Stud Constr Mater 17:e01543. https://doi.org/10.1016/J.CSCM.2022.E01543
27. Al-Ghamdi H, Elsafi M, Sayyed MI, Almuqrin AH, Tamayo P (2022) Performance of newly developed concretes incorporating WO_3 and barite as radiation shielding material. J Mater Res Technol 19:4103–4114. https://doi.org/10.1016/J.JMRT.2022.06.145
28. Lardhi M, Mukhtar F (2023) Radiation shielding performance of seawater-mixed concrete incorporating recycled coarse aggregate and steel slag. J Radiat Res Appl Sci 16:100528. https://doi.org/10.1016/J.JRRAS.2023.100528

Development of a Non-Invasive Approach to Glucose, Cholesterol, and Hemoglobin Monitoring

Nur Hasanah Ahniar[1(✉)], Siddiq Andrianto[1], Rinda Nur Hidayati[1], Gita Rindang Lestari[1], and Anita Miftahul Maghfiroh[2]

[1] Department of Electromedical Engineering, Health Polytechnic Ministry of Health, Jakarta, Indonesia
ahniar@poltekkesjkt2.ac.id

[2] Department of Electromedical Engineering, Health Polytechnic Ministry of Health, Surabaya, Indonesia

Abstract. Non-Communicable Diseases (NCDs) are a significant global health issue, responsible for 41 million deaths annually, or 74% of global mortality. Diabetes alone caused 2 million deaths in 2022, and Indonesia ranks 5th in diabetes prevalence worldwide. Cases in the country are projected to rise from 19.5 million to 28.6 million by 2045. Cardiovascular diseases also present a substantial burden, contributing to approximately 650 million deaths. Key risk factors like high systolic blood pressure and elevated fasting plasma glucose have worsened these outcomes. In 2019, diabetes mellitus accounted for 1.8% of Disability-Adjusted Life Years (DALYs) in Indonesia, emphasizing the need for effective monitoring tools. Traditional invasive methods for measuring glucose, cholesterol, and hemoglobin levels often cause discomfort, highlighting the need for a non-invasive alternative. This study aims to develop a non-invasive device using near-infrared (NIR) light at 940 nm and BPW34 photodiodes, integrated with the COCOBIN mobile application, to measure glucose, cholesterol, and hemoglobin levels. Involving 60 participants, the study compared COCOBIN device readings against a standard Glucose, Cholesterol, and Uric Acid (GCU) meter. Results indicated the COCOBIN device achieved 86% accuracy for glucose, with an error rate of 14%; 87% accuracy for cholesterol, with a 13% error rate; and 88% accuracy for hemoglobin, with a 12% error rate. These findings demonstrate the device's potential as an accurate, non-invasive alternative to traditional methods, enhancing patient compliance and supporting early health risk detection.

Keywords: Non-invasive · glucose · cholesterol · urid acid

1 Introduction

Chronic diseases such as diabetes, coronary heart disease, and anemia pose significant challenges to public health worldwide. According to the latest data from the WHO, approximately 422 million people globally suffer from diabetes, with prevalence continuing to rise each year [1]. Diabetes not only affects individuals' quality of life but

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 459–473, 2026.
https://doi.org/10.1007/978-981-95-9233-3_33

also increases the risk of serious complications, including heart disease. Coronary heart disease, closely related to risk factors such as hypercholesterolemia and diabetes, is a leading cause of death in many countries. Data from the American Heart Association indicates that 18.2 million people in the U.S. had coronary heart disease in 2021 [2]. On the other hand, anemia, often caused by iron deficiency or chronic disorders, is also a global health issue. A report from UNICEF notes that around 1.62 billion people worldwide are affected by anemia, with the highest prevalence found in developing countries [3]. These three conditions demand effective and ongoing monitoring for optimal management. Non-invasive methods for measuring glucose, cholesterol, and hemoglobin offer a promising solution to minimize patient discomfort and improve compliance with routine testing. Non-invasive devices can facilitate early detection and management of chronic diseases, which is crucial given the high costs and risks associated with disease complications. Recent research shows that technologies such as infrared spectroscopy and optical scanning have the potential to replace conventional invasive and painful methods [4].

However, challenges in the validation and accuracy of these non-invasive measurement devices remain a focus of research. Given the high incidence of diabetes, coronary heart disease, and anemia, innovations in more accessible and comfortable measurement methods should be a priority in the healthcare field.

Traditional monitoring of glucose, cholesterol, and hemoglobin levels involves invasive techniques that require blood samples, which can cause pain and deter patients from undergoing regular testing. Non-invasive devices aim to circumvent these barriers by offering a more comfortable and efficient method. Technologies like infrared (IR) spectroscopy, optical scanning, and photoplethysmography (PPG) have shown potential as alternatives to conventional blood-based measurements [5]. Among these, the BPW34 photodiode sensor, commonly used in NIR applications, has garnered significant attention for its ability to detect light intensity changes in biological tissues, which can be correlated with glucose, cholesterol, and hemoglobin concentrations.

Research on non-invasive glucose, cholesterol, and hemoglobin measurement devices over the past five years has heavily focused on the use of Near Infrared (NIR) sensors, which have the capability to detect biomarker changes through light interaction with body tissues. NIR sensors for glucose measurement have been developed using various methods, such as NIR spectroscopy and NIR-based photoplethysmography (PPG). Studies have shown that combining NIR sensors with machine learning algorithms, such as multivariate regression and neural networks, significantly improves glucose measurement accuracy [6–8]. NIR-based PPG technology has also successfully monitored glucose levels continuously and in real-time with high sensitivity, making it a promising non-invasive alternative for diabetes patients [9]. The non-invasive measurement of glucose has been one of the primary research focuses due to the global burden of diabetes. The BPW34 photodiode, when integrated with NIR spectroscopy techniques, has shown promising results in detecting glucose levels by analyzing the absorption and reflection of light through tissue. NIR spectroscopy coupled with photodiodes could achieve accurate glucose measurements by analyzing the spectral characteristics of glucose [10, 11]. By incorporating machine learning techniques, such as multivariate regression and neural networks, the accuracy of glucose measurements has been significantly improved

[12, 13]. NIR-based PPG technology, which measures the variations in light intensity as it passes through or is reflected from tissues, has also proven effective in continuously monitoring glucose levels. This approach has demonstrated high sensitivity in real-time glucose monitoring, making it a practical solution for non-invasive diabetes management [14]. Ongoing research on BPW34 photodiode applications, highlights the potential of this sensor to be integrated into portable glucose monitors, reducing the need for blood sampling [14].

In addition, research on non-invasive cholesterol measurement has progressed rapidly, with NIR sensors used to detect lipids beneath the skin. Several studies indicate that NIR, combined with reflectance spectroscopy techniques, provides accurate cholesterol measurements without requiring blood samples [9]. NIR sensors have been applied in clinical settings, demonstrating great potential for routine health monitoring. Several studies have indicated that NIR reflectance spectroscopy, combined with photodiodes, can accurately detect lipid profiles and cholesterol concentrations by analyzing the optical properties of subcutaneous tissue [15, 16], NIR spectroscopy could measure cholesterol concentrations with high accuracy, reducing the need for blood tests [17]. Application of BPW34 photodiodes in portable cholesterol measurement devices, where the combination of NIR spectroscopy and advanced signal processing techniques provided reliable readings [18].

Non-invasive hemoglobin measurement using NIR sensors has also become an intriguing research topic. Several studies leverage NIR to detect hemoglobin levels in the blood with high accuracy, particularly in daily health monitoring applications. Transdermal NIR spectroscopy studies have shown consistent results in hemoglobin measurement, making this technology an ideal non-invasive solution for both clinical and home monitoring [6, 9]. The recent advancements in NIR sensor technology are expected to drive the implementation of non-invasive measurement tools for glucose, cholesterol, and hemoglobin, reducing the need for invasive procedures [7, 9]. BPW34 sensor was used to detect variations in hemoglobin concentration based on light absorption, providing consistent results in non-invasive hemoglobin monitoring, combination of NIR technology and BPW34 photodiodes could be effectively applied in clinical settings for continuous hemoglobin monitoring, improving patient comfort and reducing the need for frequent blood draws [19, 20].

Overall, advances in NIR sensor technology have opened up significant opportunities for the development of more comfortable and efficient non-invasive medical devices. The combination of NIR sensors with machine learning data processing continues to yield promising results, which is anticipated to improve the quality of life for patients with conditions like diabetes and cardiovascular diseases [6, 7, 21].

2 Materials and Method

2.1 Experimental Setup

This study involved 75 participants aged between 20 and 30 years. This study has received ethical approval from the Health Research Ethics Committee of Politeknik Kesehatan Kemenkes Jakarta II with the number DP.04.03/I/KE/38/123/2024. The research

was conducted within the environment of Poltekkes Kemenkes Jakarta II. The respondents were randomly selected, regardless of gender or health status. In accordance with the ethical approval, each respondent received compensation in the form of food and beverages.

In this data collection, the author will conduct several measurements and comparisons. The comparison involves comparing the module with the GCU Meter, using the parameters of Blood Sugar, Cholesterol, and Hemoglobin. The author will conduct testing and data analysis for 3 measurements. Each measurement will include 30 samples. The accuracy of the data collection can be influenced by several factors, such as the assembled components and the supporting tools used to gather the data.

The data collection was carried out in two stages. The first stage involved comparing the ADC output results with the glucose, cholesterol, and hemoglobin levels obtained through invasive methods. The data were then subjected to linear regression to achieve a more stable distribution of measurement values. After obtaining the ADC linear regression values, the formula was entered into the Arduino. The second stage of data collection was conducted by comparing the values from the COCOBIN module with invasive measurement tools.

2.2 Materials and Tools

A suitability test was conducted to establish concordance between the ADC value generated by the probe and actual glucose and cholesterol levels. In this procedure, glucose and cholesterol levels were measured using the GCU Meter and the designed instrument at approximately the same time. The researcher employed an invasive glucose and cholesterol measuring device to measure the subject's glucose levels. Subsequently, to acquire the ADC value, which represents the subject's glucose and cholesterol levels, measurements were performed by placing a finger on the probe using a module developed by the researcher. This measurement was executed 50 times to determine the relationship between the ADC value on the module and the glucose and cholesterol levels on the GCU Meter. From the data obtained, an equation was formulated that represents the relationship between glucose and cholesterol levels and the ADC value. The relationship between the ADC value and blood glucose levels can be approximated by the Eq. (1) and Fig. 1:

$$y = 0.0592x + 63{,}225 \tag{1}$$

Equation (1) explains that Y is the dependent variable (the one we want to predict or explain). X is the independent variable (used to make the prediction). 0.0592 is the regression coefficient, indicating the average change in Y for each unit change in X. 63.225 is the intercept (constant), representing the value of Y when X is zero.

1. **Regression Coefficient (0.0592)**: This shows that for every one-unit increase in the variable **X**, the variable **Y** increases by 0.0592 units. This coefficient indicates the strength and direction of the relationship between **X** and **Y**.
2. **Intercept (63.225)**: This is the predicted value of **Y** when **X** is 0. It represents the point where the regression line crosses the Y-axis.

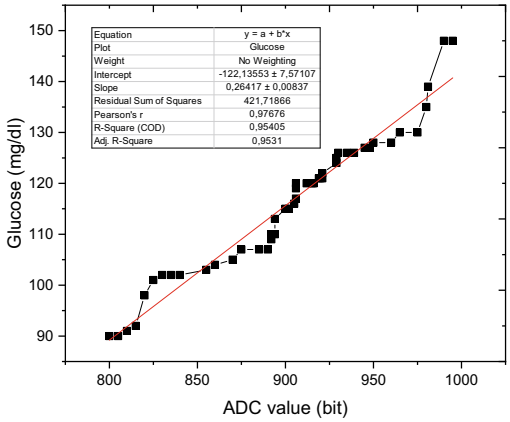


Fig. 1. The linearity between ADC and glucose values

Figure 1 shows that the linearity between ADC and glucose values. The linearity of Fig. 1 was obtained by analysis using the Origin software namely the R^2 value shows the proportion of variability in the dependent data that can be explained by the regression model. R^2 values range from 0 to 1. The closer it is to 1, the better the model explains the variability in the dependent data. An R^2 coefficient value of 0.4356 indicates that X influences Y. This equation is then incorporated into the microcontroller program as the equation that will convert the ADC value to the blood glucose level. The relationship between the ADC value and cholesterol level can be approximated by the Eq. (2) and Fig. 2:

$$y = 0.03085x + 106.06 \tag{2}$$

The correlation coefficient R = 0.3169 indicates that the value of x influences y

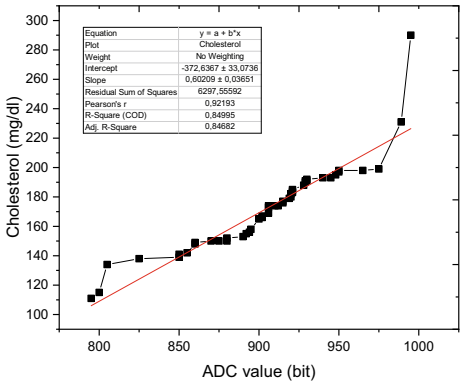


Fig. 2. The linearity between ADC and cholesterol values.

This equation is then incorporated into the microcontroller program as the equation that will convert the ADC value into the cholesterol level. The R^2 value shows how much the independent variable (exogenous) affects the dependent variable (endogenous). Based on Eq. (3) and Fig. 3, the relationship between the ADC value and hemoglobin level can be approximated by the equation:

$$y = 0.0123x + 2.756 \tag{3}$$

The correlation coefficient $R = 0.459$ indicates that x influences y. This equation is then incorporated into the microcontroller program as the equation that will convert the ADC value into the cholesterol level. The R^2 value shows how much the independent variable (exogenous) affects the dependent variable (endogenous). This data was collected in a single session using the Cocobin device and the GCU meter. In this study, a total of 50 data samples were carefully collected to ensure the representativeness and reliability of the results. The data collection was conducted with attention to the variations present in the studied population, ensuring that the samples taken could accurately represent the actual conditions. After the data was collected, the next step was to perform an analysis using the linear regression method.

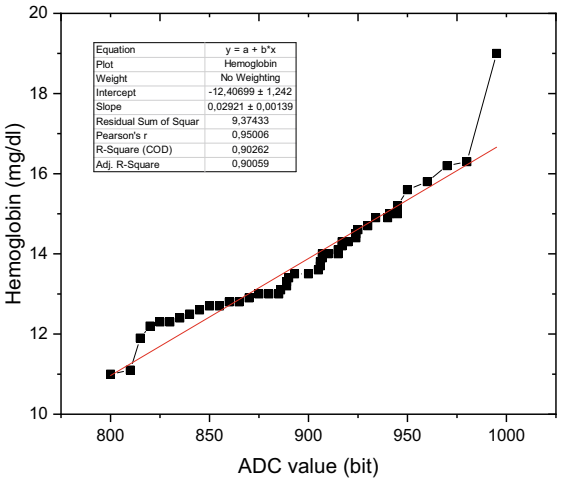


Fig. 3. The linearity between ADC and hemoglobin values.

The linear regression method was chosen for its ability to identify and measure the relationship between the independent and dependent variables of interest in this research. By applying linear regression, the researchers can test the proposed hypothesis and evaluate the strength and direction of the relationship between the variables. This analysis also provides deeper insights into the patterns and trends present in the data, making the results not only relevant but also useful as a basis for further decision-making. The findings from this analysis are expected to contribute significantly to understanding the dynamics between the variables studied.

2.3 Block Diagrams

Figure 4 is a block diagram of the module design used. In the block diagram there are three parts, namely: input, process, and output. The Supply Block functions to provide voltage to the entire circuit and the components used. The LED Block serves as the light source on the probe, which will pass through the patient's finger and be captured by the photodiode. The Photodiode Block functions to capture the light from the LED that has passed through the patient's finger. The intensity of the captured light affects the resistance of the photodiode, which influences the amount of electrical current flowing. The Microcontroller Block serves as the control center of the device and processes the data obtained from the photodiode sensor's measurements. The microcontroller used in this case is the NodeMCU ESP32. The LCD Block functions as the display of the device, showing the measurement results and instructions on how to operate the device.

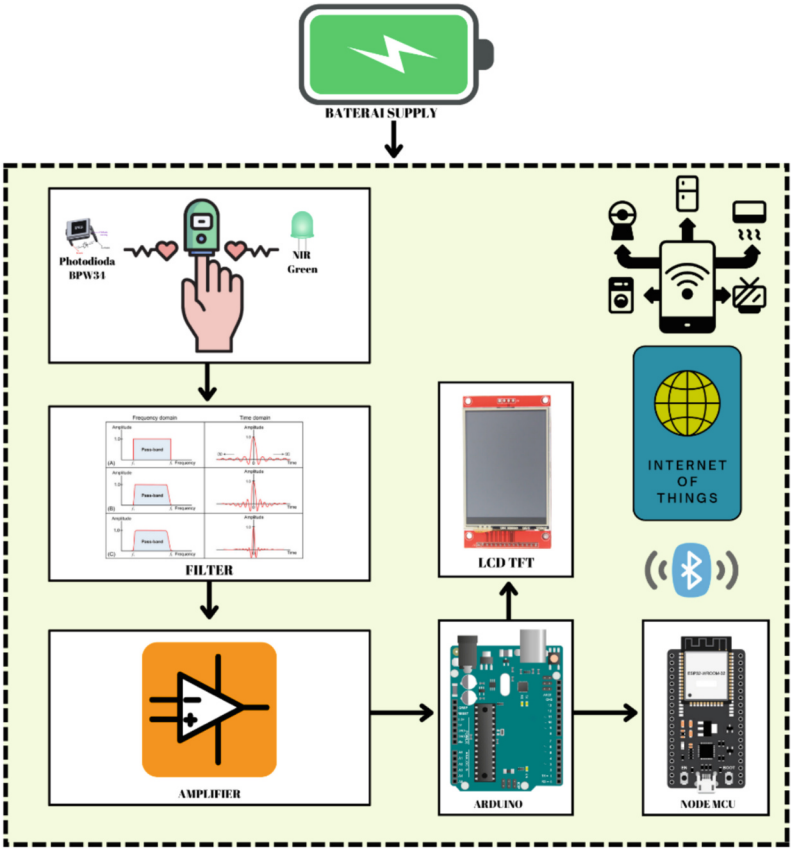


Fig. 4. Block diagram of the cocobin module

In the supply block, there is a lithium battery, which is a rechargeable battery. This lithium battery is connected to a battery management IC, which is responsible for recharging the battery. The voltage from the battery goes to a boost converter, where the voltage is increased from 3.7 V to the required 5 V. Next, in the input block, there is a photodiode that functions as a sensor. Infrared and green LEDs serve as light sources. The output then goes to the ADS1115, which functions as an ADC and amplifier, converting analog data into digital data and amplifying the input signal. This data is then processed by the NodeMCU ESP32 and sent to the TFT LCD and also to an Android device.

3 Result

The output voltage was assessed by measuring the voltage across the resistor configured as a voltage divider with a photodiode. This voltage value was subsequently converted into an Analog-to-Digital Converter (ADC) value and processed by the microcontroller to determine glucose, Cholesterol, and hemoglobin levels. The measurements were conducted under two conditions: without a finger on the probe and with a finger on the probe as shown as Table 1.

Table 1. Output voltage measurement

Conditions	V_{out}
without a finger	0.29 V
with a finger	0.18 V

The measurements were conducted using a digital multimeter by positioning the black probe on the ground and the red probe on the resistor leg connected to the anode leg of the photodiode. The voltage in the condition without a finger exceeded that in the condition with a finger. This phenomenon occurs because, in the absence of a finger, light is reflected in the probe without any absorption of light intensity. Conversely, when a finger is present, it absorbs and reflects a portion of the light, resulting in a lower light intensity captured by the photodiode compared with the condition without a finger. It can be concluded that in this apparatus, the light intensity received by the photodiode is directly proportional to the output.

3.1 Evaluate the Functionality of the Device

To ensure the accuracy of the blood sugar, cholesterol and hemoglobin monitoring function test, we conducted a series of meticulous data-collection procedures. Data collection was performed 60 times, with each iteration comprising 5 consecutive trials. The purpose of this repetition was to minimize the potential for errors or anomalies in measurement results. Each measurement was executed with precision, ensuring that the conditions of the device and subject remained consistent throughout the data collection process. This consistency was crucial for ensuring that the results accurately reflected the actual

conditions and were not influenced by uncontrolled external factors. The results from each of the five trials were carefully recorded. Following data collection, we calculated the mean of the five measurement values obtained to derive a representative value for each data collection instance. This was considered the final result of the data collection. The process of collecting, recording, and calculating the mean was repeated consistently until 60 sets of data were accumulated for further analysis. The data generated from this process were subsequently utilized to conduct a comprehensive analysis to assess the reliability and accuracy of the device for monitoring blood sugar levels.

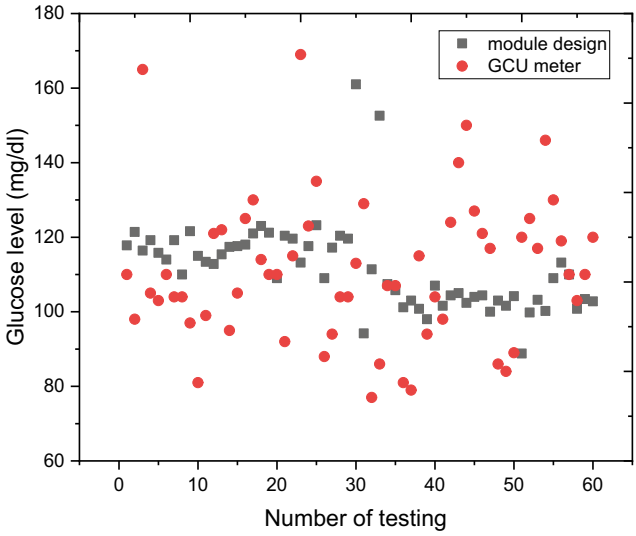


Fig. 5. The number of testings for glucose levels

In Fig. 5, the y-axis represents the blood sugar level, while the x-axis depicts the measurement sequence from the 1st sample to the 60th sample. This measurement compares glucose levels obtained using cocobin (module) and GCU Meter. The graph illustrates the variation in blood sugar levels for each measurement from two methods, namely Cocobin and GCU meter, demonstrating the fluctuation of the measured blood sugar levels. Factors influencing these levels include insulin production, insulin sensitivity, hormones, and medical conditions. The data collection method for this study involved utilizing 60 samples and comparing blood sugar levels measured with a Cocobin and GCU meter. Based on the aforementioned calculations, it can be concluded that the percentage accuracy of the blood sugar measurement module was 86%, with an average error percentage of 14%. The factors contributing to the error value include sensor placement, movement during examination, LED sensitivity, skin thickness and pigmentation, and ambient light.

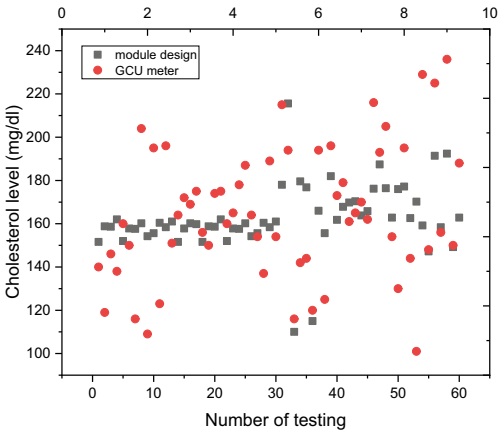


Fig. 6. The number of testings for cholesterol levels

The Figs. 6 and 7 depicts the variation in cholesterol and hemoglobin levels for each measurement obtained from two methods, namely Cocobin and GCU Meter, illustrating the fluctuation of measured cholesterol levels. This fluctuation may be influenced by various factors including diet, physical activity, medications, and other health conditions. The methodology for data collection in this study involved the use of 60 samples, comparing cholesterol measurements obtained from Cocobin and GCU meters. Based on the aforementioned calculations for cholesterol levels in Fig. 6, it can be concluded that the percentage accuracy of the cholesterol measurement module is 87%, with an average error percentage of 13%. Based on the results of the hemoglobin test presented in Fig. 7, it can be concluded that the hemoglobin measurement module demonstrated an accuracy percentage of 88% with a mean error percentage of 12%.

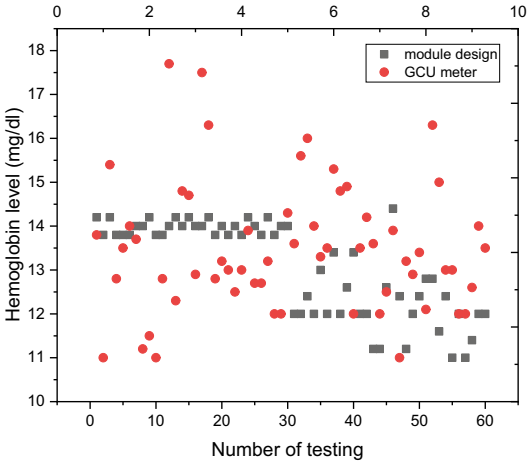


Fig. 7. The number of testings for hemoglobin levels

3.2 Battery Functional Testing

This section elucidates the battery-testing methodology employed as a power source for the designed tool. The objective of battery testing is to determine the duration for which the battery can sustain power to the tool. This examination was conducted using three consecutive tests. Each test aimed to compare the duration recorded on the stopwatch until the tool ceased operation. The testing process was executed meticulously to ensure accurate and consistent results. Furthermore, an analysis of the test results will provide a comprehensive understanding of the tool's performance under the tested conditions, as well as aid in evaluating its reliability for long-term use. In Table 2, the first ten trials involved charging the cocobin tool battery from empty (0%) to full (100%) to assess the duration required for a complete charge. The second trial was conducted to determine the duration of the standby (on) condition until battery depletion. The third trial was performed to ascertain the operational duration of the tool when it was utilized by the user until battery exhaustion.

Table 2. Battery functional testing

No	Batteray capacity	Time (hour)
1	Consumed—Full	8
2	Full—Consumed (Device Standby)	6
3	Full—Consumed (Device Used)	5

4 Discussion

The research on the development of a non-invasive device for measuring glucose, cholesterol, and hemoglobin shows significant advantages, but also reveals several weaknesses and limitations that need to be addressed. This device is designed to measure these three important biomarkers in the blood without the need for direct blood sampling, utilizing optical sensor technology and photodiodes. Although the measurement results indicate good potential, there are several aspects that require further discussion, especially concerning accuracy, device stability, comparisons with invasive methods, as well as similar research, and the potential environmental impact.

One of the main weaknesses of this non-invasive device is its sensitivity to environmental and individual conditions. Measurements of glucose, cholesterol, and hemoglobin, which rely on light reflected or absorbed by the skin and tissue, can be influenced by various factors such as skin thickness, pigmentation, temperature, and the patient's hydration levels. In trials, the device showed significant variability in results among subjects with different physiological characteristics, often requiring recalibration. This leads to a decrease in accuracy compared to invasive methods that directly measure biomarker levels in the blood.

Another limitation is the data processing and analysis algorithms used. Although technologies such as the NodeMCU ESP32 microcontroller and photodiode sensors

used can provide fast data processing, the current data processing algorithms still need improvement to more effectively filter out noise caused by environmental interference or variations in optical signals. This challenge suggests the need for further optimization in system design and data interpretation algorithms to produce more consistent and accurate results.

Compared to previous studies, this device offers a more practical and user-friendly approach, as it eliminates the need for skin puncture or blood sampling. Some earlier studies employed techniques like near-infrared spectroscopy (NIR) and Raman spectroscopy for non-invasive measurements. However, these approaches often require larger, more expensive equipment and are difficult to apply in portable devices for daily use. In this regard, the device developed in this research is more portable and potentially more cost-effective for widespread use.

However, other research, such as studies using photoplethysmography, has demonstrated higher accuracy in non-invasive hemoglobin measurements. Therefore, integrating this approach with the photodiode sensors used in this device might be a solution for improving measurement outcomes. Additionally, other studies using AI and machine learning to process data from optical sensors could serve as a reference for enhancing the signal processing algorithms of this device.

The use of this non-invasive device could have a positive impact on both the environment and society. Environmentally, this device reduces medical waste generated from disposable lancets, needles, and blood test strips typically required in invasive methods. By reducing reliance on conventional testing tools, the amount of hazardous medical waste can be significantly minimized. Additionally, by reducing the need for chemical reagents used in laboratory analyses, this device contributes to lowering chemical pollution.

The social impact of this device is also significant. It could increase accessibility for patients who require regular monitoring of glucose, cholesterol, and hemoglobin, especially in areas far from healthcare facilities. The availability of a practical and non-invasive device like this could improve the quality of life for patients suffering from chronic conditions such as diabetes and hypercholesterolemia. Moreover, by enabling self-monitoring of health, patients would not need to visit healthcare facilities as frequently, thus reducing the burden on the healthcare system.

Although this non-invasive measurement device holds great potential to revolutionize the way glucose, cholesterol, and hemoglobin levels are measured, several challenges still need to be addressed, particularly in terms of measurement accuracy and result variability. The integration of new technologies and improved signal processing algorithms will be key to achieving better performance. The positive environmental and social impacts of using this device also highlight its potential as an important tool in the future of healthcare technology development.

This study shows that non-invasive tools tend to have significant variability in results among individuals with different physiological characteristics. This is in contrast to previous studies using near infrared (NIR) spectroscopy and Raman spectroscopy, which can provide higher accuracy, although they often require larger and more expensive devices [22, 23]. The developed tools are still inferior in accuracy compared to invasive methods, which measure biomarkers directly in the blood. For example, research by

Wong et al. (2019) [23] shows that invasive methods remain the gold standard for measuring glucose levels. Another study using photoplethysmography (PPG) showed higher accuracy in non-invasive hemoglobin measurement. The integration of PPG technology with a photodiode sensor in this study can improve measurement results [24]. Limitations in current processing algorithms hamper the accuracy of the tool. Other studies using AI and machine learning for optical sensor data analysis have shown that this approach can reduce noise resulting from environmental variations [25]. The results of this study indicate lower accuracy compared to previous research due to the sensor design, which utilizes a single sensor for the detection of haemoglobin, glucose, and cholesterol parameters. Each parameter may have different physical or chemical characteristics, making it difficult for a single sensor to detect all with the same level of precision. In the use of a single sensor, the signal generated by one parameter can interfere with the signals from other parameters. This may result in biased or inaccurate outcomes, as the sensor may struggle to differentiate the specific response of each parameter [26, 27].

5 Conclusions

The design of the non-invasive instrument for detecting glucose, cholesterol, and hemoglobin levels has been successfully developed, showing promising results in its initial testing phase. The device demonstrated an average accuracy of 86% in measuring blood sugar levels, with an error margin of 14%. For cholesterol levels, the device exhibited an average accuracy of 87%, with an error margin of 13%. Similarly, in measuring hemoglobin levels, it achieved an average accuracy of 87%, with an error margin of 13%. This performance underscores the potential of non-invasive monitoring devices to serve as reliable tools for health management, particularly for individuals requiring regular monitoring of these key biomarkers.

One of the primary advantages of this module is its ability to perform measurements without the need for blood samples, effectively eliminating the need for invasive procedures such as finger pricking, lancet use, and test strips. This not only enhances the convenience and comfort of the user but also significantly reduces the generation of medical waste associated with single-use test kits, aligning the design with sustainable healthcare practices. As the world moves toward environmentally conscious solutions, this device represents a step forward in reducing healthcare-related environmental impact.

While the current design has yielded favorable results, the accuracy levels, although acceptable, still leave room for improvement. An error margin of 13–14% can result in some variability in clinical decision-making, especially for individuals with critical glucose, cholesterol, or hemoglobin levels. Given that the device's accuracy is affected by physiological factors such as skin thickness, hydration, and tissue composition, future refinement of the technology should focus on minimizing the impact of these variables.

Furthermore, the current design's reliance on optical sensors and photodiodes, while effective, could be further enhanced by integrating more advanced signal processing techniques. The observed deviations in accuracy are largely attributed to external interference and noise within the sensor readings. To address this, future iterations could incorporate machine learning algorithms to better interpret the data collected from the sensors and adjust for individual variability, thereby enhancing the precision of the device.

While the non-invasive glucose, cholesterol, and hemoglobin measurement device has demonstrated significant potential in providing an accessible, user-friendly, and environmentally conscious solution, continued research and development are essential for enhancing its accuracy and utility. With a focus on advanced data processing techniques, clinical validation, miniaturization, and interconnectivity, this device could revolutionize the way individuals manage their health on a daily basis, offering a practical alternative to traditional invasive methods. Additionally, by reducing medical waste and contributing to sustainable healthcare, this innovation aligns with broader goals for environmental stewardship and responsible healthcare practices.

References

1. World Health Organization: Diabetes. https://www.who.int/news-room/fact-sheets/detail/diabetes. Last Accessed 28 Oct 2024
2. American Heart Association: Heart Disease and Stroke Statistics. https://www.heart.org/en/about-us/heart-and-stroke-association-statistics, Last Accessed 28 Oct 2024
3. UNICEF. The state of the world's children. 2019, https://www.unicef.org/reports/state-worlds-children-2019, Last Accessed 28 Oct 2024
4. Lin T (2017) Non-invasive glucose monitoring: a review of challenges and recent advances. Curr Trends Biomed Eng Biosci 6:1–8. https://doi.org/10.19080/CTBEB.2017.06.555696
5. Alsunaidi B, Althobaiti M, Tamal M, Albaker W, Al-Naib I (2021) A review of non-invasive optical systems for continuous blood glucose monitoring. Sensors 21:6820. https://doi.org/10.3390/s21206820
6. Hina A, Saadeh W (2022) Noninvasive blood glucose monitoring systems using near-infrared technology—a review. Sensors 22:4855. https://doi.org/10.3390/s22134855
7. Valero M, Ingram K, Duong A, Nino V (2024) Pervasive glucose monitoring: a non-invasive approach based on near-infrared spectroscopy. In: Salvi D, Van Gorp P, Shah SA (eds) Pervasive computing Technologies for Healthcare. Springer Nature Switzerland, Cham, pp 274–289
8. González-Legarreta L et al (2015) Sensing technology: current status and future trends IV
9. Ramasahayam S, Sri Haindavi K, Chowdhury SR (2015) Noninvasive estimation of blood glucose concentration using near infrared Optodes. In: Mason A, Mukhopadhyay SC, Jayasundera KP (eds) Sensing technology: current status and future trends IV. Springer International Publishing, Cham, pp 67–82. https://doi.org/10.1007/978-3-319-12898-6_4
10. Wang Q, Wu G, Pian F, Shan P, Li Z, Ma Z (2021) Simultaneous detection of glucose, triglycerides, and total cholesterol in whole blood by Fourier-transform Raman spectroscopy. Spectrochim Acta A Mol Biomol Spectrosc 260:119906. https://doi.org/10.1016/j.saa.2021.119906
11. Reddy PSK, Mahesh D, Uday Teja C, Janaki M, Mannem K (2022) Non-invasive glucose monitoring using NIR spectroscopy. J Phys Conf Ser 2325:012021. https://doi.org/10.1088/1742-6596/2325/1/012021
12. Manurung BE, Munggaran HR, Ramadhan GF, Koesoema AP (2019) Non-invasive blood glucose monitoring using near-infrared spectroscopy based on internet of things using machine learning. In: 2019 IEEE R10 Humanitarian technology conference (R10-HTC) (47129), pp 5–11. https://doi.org/10.1109/R10-HTC47129.2019.9042479
13. Agrawal H, Jain P, Joshi AM (2022) Machine learning models for non-invasive glucose measurement: towards diabetes management in smart healthcare. Health Technol (Berl) 12:955–970. https://doi.org/10.1007/s12553-022-00690-7

14. Daniyan I, Ikuponiyi S, Daniyan L, Damian Uchegbu I, Mpofu K (2022) Development of a smart glucose monitoring device. Procedia CIRP 110:253–258. https://doi.org/10.1016/j.procir.2022.06.046
15. K R. (2021) NIR reflectance spectroscopy for cholesterol detection: a non-invasive approach. J Clin Med 10:101–115
16. Anupongongarch P, Kaewgun T, Ananpatiwet P (2020) A study on design and construction of non-invasive cholesterol sensor. In: RSU international research conference 2020, pp 1–10
17. Umar U, Syarif S, Nurtanio I (2020) Indrabayu: a real time non-invasive cholesterol monitoring system. In: MATEC web of conferences, vol 331, pp 06005. https://doi.org/10.1051/matecconf/202033106005
18. Pavithra K-S, Mary XA, Rajasekaran K, Jegan R (2017) Low cost non-invasive medical device for measuring hemoglobin. In: 2017 International conference on innovations in electrical, electronics, instrumentation and media technology (ICEEIMT), pp 197–200. https://doi.org/10.1109/ICIEEIMT.2017.8116834
19. Nandanwar R, Sayre S, Sameen S, Ramteke S, Jaiswal S, Kasetwar A (2021) Non-invasive sensor Technology for Hemoglobin Measurement in blood. Int J Sci Res Eng Manag (IJSREM) 5:1–5
20. Rochmanto RA, Zakaria H, Alviana RD, Shahib N (2017) Non-invasive hemoglobin measurement for Anemia diagnosis. In: 2017 4th International conference on electrical engineering, computer science and informatics (EECSI), pp 1–5. https://doi.org/10.1109/EECSI.2017.8239096
21. Zafar H, Channa A, Jeoti V, Stojanović GM (2022) Comprehensive review on wearable sweat-glucose sensors for continuous glucose monitoring. Sensors 22:638. https://doi.org/10.3390/s22020638
22. Devakumar VJ, Ravi T, Karthikeyan S (2019) Review on non-invasive glucose and cholesterol measurement system. IOP Conf Ser Mater Sci Eng 590:012030. https://doi.org/10.1088/1757-899X/590/1/012030
23. Ahmadian N, Manickavasagan A, Ali A (2023) Comparative assessment of blood glucose monitoring techniques: a review. J Med Eng Technol 47:121–130. https://doi.org/10.1080/03091902.2022.2100496
24. Lychagov VV, Semenov VM, Volkova EK, Chernakov DI, Ahn J, Kim JY (2023) Noninvasive Hemoglobin measurements with Photoplethysmography in wrist. IEEE Access 11:79636–79647. https://doi.org/10.1109/ACCESS.2023.3300293
25. Gupta S, Bhatt PK, Mishra S, Kumar S (2025) Artificial intelligence applications in optical sensor technology. In: Jaiswal A, Anand S, Hassanien AE, Azar AT (eds) Proceedings of international conference on computing and communication Systems for Industrial Applications. Springer Nature Singapore, Singapore, pp 179–193
26. Awad AA-U (2018) Environmental impacts of medical waste treatment and management by burning inside health facilities. Int J Civ Eng Technol (IJCIET) 9:41–53
27. Ates HC et al (2021) Integrated devices for non-invasive diagnostics. Adv Funct Mater 31:1–18. https://doi.org/10.1002/adfm.202010388

Detecting Major Depressive Disorder or Depression in Electroencephalogram Signal Using Long Short-Term Memory Networks

Mera Kartika Delimayanti[1(✉)], Euis Oktavianti[1], Bambang Warsuta[1], Rizky Adi[1], Umar Ali Ahmad[2], and Kenji Satou[3]

[1] Computer and Informatics Engineering Department, Politeknik Negeri Jakarta, Depok, Indonesia
mera.kartika@tik.pnj.ac.id
[2] School of Electrical Engineering, Telkom University, Bandung, Indonesia
[3] Institute of Science and Engineering, Kanazawa University, Kanazawa, Japan

Abstract. Major Depressive Disorder (MDD) is a prevalent mental health condition with limited treatment efficacy. Electroencephalography (EEG) offers a promising avenue for exploring the neural underpinnings of MDD. However, existing EEG-based approaches often rely on time-consuming feature extraction methods and require a large number of recording electrodes. This can be impractical for clinical settings. This study proposes a novel approach using Long Short-Term Memory (LSTM) networks to classify MDD based on 128-channel EEG resting state recordings. The dataset used in this research is MODMA (Multi-model Open Dataset for Mental-disorder Analysis). The LSTM used in the research is a type of recurrent neural network, that can directly learn from raw EEG data, eliminating the need for manual feature engineering. The results demonstrate an accuracy of 72% in distinguishing individuals with MDD from healthy controls. This research suggests the potential of LSTM for accurate and efficient MDD classification using EEG data, paving the way for improved diagnosis and treatment.

Keywords: Deep Learning · EEG · LSTM · Major Depressive Disorder · MODMA

1 Introduction

Major Depressive Disorder (MDD), commonly known as depression, is a prevalent and severe psychiatric disorder globally, which in extreme cases can lead to suicide [1–3]. This mental health condition is characterized by persistent sadness, poor concentration, low energy levels, disrupted sleep and changes in appetite or weight [1, 4]. The World Health Organization estimated that 5% of adults suffer from depression worldwide [4, 5]. Not only that, depression also can affect adolescence and teenager [6–8]. Thus, showing the importance of early detection of depression to prevent extreme cases such as suicides [2, 6]. The earlier the depression is detected, the easier the treatment can be conducted.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 474–484, 2026.
https://doi.org/10.1007/978-981-95-9233-3_34

Currently, the process to diagnose depression is conducted through a scale-based interview by a psychologist or psychiatrist. The international standards that are mostly used are Mini-Mental State Examination and DSM-IV [9, 10]. Even though the standard used currently already great, it needs a professional analysis to do the interview for each patient. Which makes the scaled-based interview diagnosis less efficient. So, there is a need to find a more efficient way to detect depression without the need for a professional to diagnose depression manually.

Electroencephalography (EEG) signal has emerged as a promising tool for diagnosing various diseases [11–15]. By measuring brain electrical activity, EEG can provide insights into abnormal brain patterns associated with certain diseases. This approach also has been used to diagnose depression using EEG signal data. Previous studies have identified differences in EEG signals between individuals with MDD and healthy controls, suggesting that EEG-based features could potentially aid in diagnosis and treatment [16–18]. Even tough existing EEG-based depression detection already yielding a promising result, these approaches often rely on time-consuming feature extraction methods that need a comprehensive analysis of each electrode [19]. This can be impractical if there are changes to clinical settings. To address these limitations, this study proposes an approach of using deep learning, especially Long Short-Term Memory (LSTM) networks [20]. LSTM is a type of recurrent neural network that is well-suited for processing sequential data like EEG signals [21]. By directly learning from raw EEG data, LSTM can automatically extract relevant features without the need for manual feature engineering.

The primary objective of this research is to develop a robust EEG-based classification model for Major Depressive Disorder (MDD) using Long Short-Term Memory (LSTM) networks relying on raw EEG signal data. This approach deviates from conventional methods that typically rely on complex features of engineering. By directly training the LSTM model on raw EEG data, we aim to explore the LSTM model's ability to capture depression patterns within the EEG signal. This approach has the potential to streamline the classification process and reduce the need for domain-specific feature engineering expertise. To provide a comparative baseline, we also investigate the performance of a model using Fast Fourier Transform (FFT) for minimal feature extraction. This allows us to assess the relative effectiveness of LSTM in capturing relevant EEG characteristics without extensive feature engineering. Overall, the contributions of this research are developing LSTM deep learning model for depression detection using EEG signal, evaluating the performance of the model, and benchmarking our results against previous study.

2 Method

This research was conducted through multiple steps. Each step is needed to ensure the success of the research. The first step is to collect the dataset, we especially collect a publicly available EEG dataset that has been labeled. After that, we split the dataset into three parts, which are training, validation, and testing data. We ensure there is no overlapping patient in each part, ensure the integrity of the research and the dataset. After that, we apply preprocessing to the dataset, the preprocessing consists of epoching and FFT (Fast Fourier Transform) [22, 23], however the FFT is optional. And then, the

training and validation data are used for training process, while the test data is used for evaluation purposes. The workflow of the research is shown in Fig. 1.

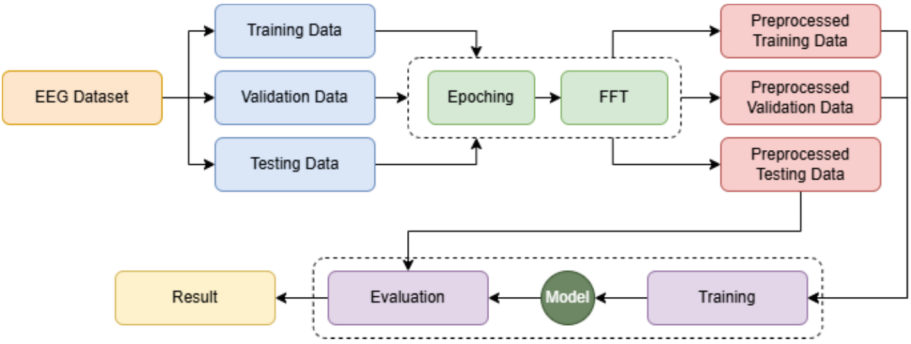


Fig. 1. Research workflow

2.1 Dataset

The EEG dataset used in this research is MODMA (Multi-model Open Dataset for Mental-disorder Analysis) [24]. Even though the dataset provides multiple dataset types namely 128-channels EEG, 3-channels EEG, and Audio dataset, we decide to only focus on 128-channels dataset as it is the most used format EEG data. The dataset was collected from 53 participants, with 24 people diagnosed with depression while the other 29 people as healthy controls. The distributions of outpatient's data are 13 males and 11 females, with age ranging from 16-to 56-year-old. While the healthy controls consist of 20 males and 9 females, with age ranging from 18-to 55-year-old. All the data captured with sampling frequency of 250 Hz for 5 min in resting state [24].

2.2 Epoching

Epoching is a crucial preprocessing step in EEG research that involves segmenting continuous EEG data into smaller, time-limited windows known as epochs [22, 25, 26]. These epochs are typically aligned with specific events or time points of interest, such as stimulus onset or response initiation. In this study, we employed a 30-s epoching window with a 15-s overlap. This configuration was chosen to balance the need for sufficient data points within each epoch to capture meaningful neural activity while also minimizing the impact of boundary effects. The 30-s epoch length provides a reasonable duration to analyze dynamic neural responses to events, while the 15-s overlap ensures a smooth transition between adjacent epochs, reducing the risk of missing important neural events. By carefully considering the rationale for epoching parameters and addressing potential limitations, we can effectively utilize this preprocessing technique to extract meaningful insights from EEG data in our research.

2.3 FFT (Fourier Fast Transform)

The Fast Fourier Transform (FFT) is a computationally efficient algorithm for computing Discrete Fourier Transforms (DFTs) [23]. DFTs are mathematical tools used to decompose periodic waveforms into their constituent harmonic components. While FFT is not ideal for analyzing short-duration brain signals, it is well-suited for processing sinusoidal signals, such as those found in electroencephalography (EEG) [27], due to its computational efficiency compared to other methods. To assess the impact of FFT on our model's performance, we conducted two sets of experiments: one using the original dataset and another using a dataset that underwent feature extraction using FFT. The formula for DFT can be seen in Eq. (1) [28].

$$(Fx)_k = \sum_{j=0}^{n-1} x_j e^{-i2\pi kj/n} \tag{1}$$

Suppose $x = (x_0, \ldots, x_{n-1})$, $w = (w_0, \ldots, w_{n-1}) \in R^n$. To evaluate F_x in need to takes $O(n^2)$ operations. However this formula implemented faster in FFT up to O(n log n) operations [23].

2.4 LSTM (Long-Short Term Memory)

The Long Short-Term Memory (LSTM) architecture offers a robust solution to the vanishing gradient problem inherent in Recurrent Neural Networks (RNNs) [29–31]. By introducing a gated memory cell, LSTM significantly enhances the ability of RNNs to learn and retain long-term dependencies within input sequences. This memory cell operates similarly to the human cognitive process of selective recall, determining whether to maintain or discard information based on its relevance. Through this mechanism, LSTM effectively extracts and preserves crucial features from the input data, enabling it to differentiate between important and less significant information [29]. The LSTM model typically comprises three gates: the forget gate f(k), the input gate i(k), and the output gate o(k) [29]. The formulas for each gate can be seen on Eq. (2), (3), and (4) respectively [32].

$$f(k) = \sigma\Big(W_f(h(k-1),\ x(k)) + b_f \tag{2}$$

$$i(k) = \sigma\Big(W_i(h(k-1),\ x(k)) + b_i \tag{3}$$

$$o(k) = \sigma\Big(W_o(h(k-1),\ x(k)) + b_o \tag{4}$$

where σ is the sigmoid activation function. The forget gate f(k) in the equation determines which part of the memory is retained for the next step. This gate produces a value between 0 and 1, and when it's 1, the entire memory cell is passed to the next state. Conversely, when it's 0, the memory is completely discarded. This mechanism allows the model to remember information for both short and long periods, hence the name "Long-Short Term Memory. These gates collaboratively manage the flow of information within

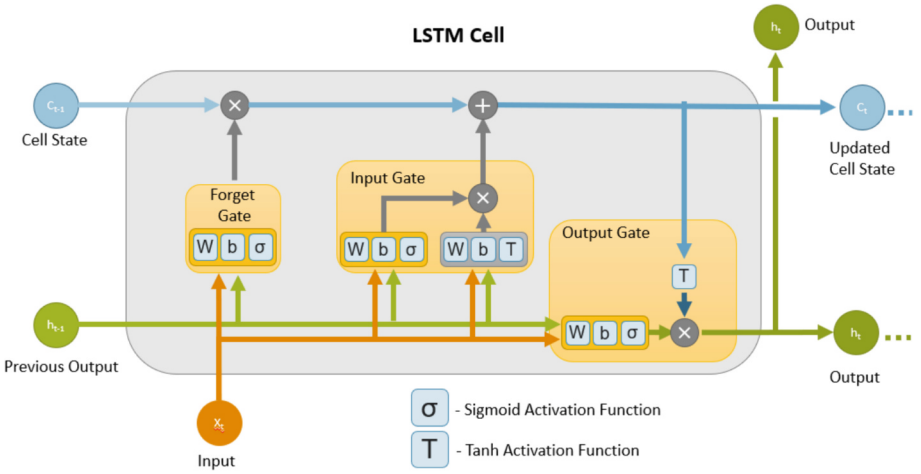


Fig. 2. LSTM cell architecture [33]

the memory cell, ensuring that only relevant data is retained, and that the model's output is influenced by the most pertinent information. The LSTM cell architecture illustration shown in Fig. 2.

The LSTM model's exceptional ability to handle long-term dependencies makes it particularly well-suited for analyzing complex time series data, such as Electroencephalogram (EEG) recordings [34, 35]. EEG data, which captures the electrical activity of the brain, often exhibits intricate patterns and temporal correlations that are difficult to model using traditional methods. LSTM can effectively learn these patterns, enabling accurate classification of EEG signals for various applications, including detecting seizures, identifying sleep stages, and diagnosing neurological disorders. By leveraging LSTM's ability to capture both short-term and long-term dependencies, the model can gain valuable insights into the dynamics of brain activity and classify the signal into distinct classes.

2.5 Evaluation

Evaluation metrics such as accuracy, precision, recall, and F1 score are commonly employed to assess the performance of trained machine learning models. Accuracy measures the overall correctness of predictions, while precision focuses on the proportion of correct positive predictions among all positive predictions. Recall gauges the ability of the model to identify all relevant positive instances. The F1 score provides a balanced measure of precision and recall, combining them into a single metric. To compute these metrics, the confusion matrix, containing true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), is utilized. Eq. (5), (6), (7, 8) present the formulas for calculating each evaluation metric.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{8}$$

3 Result

The EEG dataset was employed to train and assess the model's performance. The training and validation datasets were employed in training the model. The test data was utilized for the purpose of assessing the efficacy of the model. During the training phase, the model was trained using training data, while the validation data was employed to assess and monitor the training process to mitigate the risk of overfitting. The present work employed the early stopping technique to mitigate the overfitting risk, preventing excessive training. The training procedure will terminate automatically by implementing early stopping if substantial progress is lacking. If there is no substantial enhancement in validation loss for a continuous span of 5 epochs, the training process will be terminated automatically. A learning rate of 1e-3 or 0.001 in the training phase was established, along with a mini-batch size of 8 and a maximum of 100 training epochs. After going through multiple experiments to train LSTM architecture model to predict whether EEG signal is a MDD or healthy, we used test dataset to evaluate the said model. The evaluation result of each model is shown in Table 1.

Table 1. Evaluation results

Model	Accuracy	Recall	Precision	F1-score
LSTM—Raw	0.72	0.8	0.66	0.72
LSTM—FFT	0.54	0.4	0.5	0.44

Training depression detection model with raw EEG data produced better results than training using FFT dataset as shown in Table 1. LSTM model that trained with raw data exceeded the other model in all evaluation metrics. This means that LSTM learns more meaningful features from raw data compared to FFT data. More detailed test results can be seen on confusion matrix in Fig. 3.

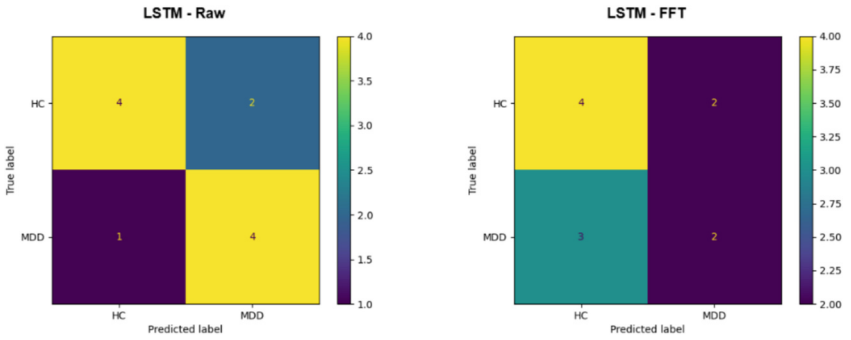


Fig. 3. Confusion matrix

Confusion matrix presented in Fig. 3 clearly shows the superiority of LSTM raw dataset model. Out of 11 test subjects, LSTM raw dataset model successfully predicted 8 patients correctly, however this model incorrectly predicted one depressed patient as healthy. Having said that, LSTM raw dataset is so much better compared to LSTM FFT dataset. The LSTM FFT dataset failed miserably to predict depressed patients, out of 5 depressed patients, LSTM FFT dataset can only predict 2 of them correctly.

4 Discussion

The result of the research shows there is a potential of training LSTM using raw electroencephalography (EEG) data for diagnosing Major Depressive Disorder (MDD). Directly learning from raw data without the need for rigorous feature engineering will shorten the development of the model a lot. This method presents itself as an alternative approach to time consuming feature engineering-based techniques providing a faster development time. Furthermore, we experimented with FFT as feature engineering to train the LSTM and the results show the model that trained on raw data has better performance. The performance of the FFT feature based models was comparatively poorer than those trained with raw data suggesting that training depression detection LSTM model is more efficient with raw data.

Table 2. Evaluation results comparison

Model	Accuracy	Recall	Precision	F1-score
CNN + GRU + Brainmap feature extraction [13]	0.89	0.90	–	0.90
CNN + Speech Data [36]	0.84	0.76	0.88	0.82
Empirical mode decomposition [37]	0.88	0.88	–	–
LSTM—Raw	0.72	0.8	0.66	0.72

Table 2 reveals the evaluation result of our best model compared to other approaches in terms of four performance metrics. Accuracy tells the proportion of correct predictions

out of the total number of predictions, recall shows the proportion of true positives out of the total number of actual positives, precision tells the proportion of true positives out of the total number of predicted positives, while F1-score harmonic mean of precision and recall, providing a balanced measure of model performance. The bigger the number shown in the table means the better is the model. And we can see that our current model produces worser result than the previous work in terms of performance. It's important to note that the other previous approaches employed a more complex feature extraction technique. Despite this, our LSTM model using raw EEG data achieved a surprisingly strong performance, nearly rivaling the previous work.

The superior performance of the brainmap-based and empirical mode decomposition methods compared to the raw EEG approach used in the previous study indicates that the additional information captured by brainmaps is crucial for accurate depression detection [16, 37]. This suggests that while raw EEG data contains a wealth of information, it can be challenging to extract meaningful features directly from it. Brainmaps, on the other hand, aggregate information from multiple EEG channels and frequency bands, potentially capturing more relevant patterns related to depression. Other works also consider using speech data to train the depression detection model [36]. The research shows a promising result with a score of 84% accuracy. However, to train that model, additional speech recording data was needed. While our approach only uses raw EEG data without feature engineering.

One of the main disadvantages of our approach is that the performance of the trained model is lower when compared to other studies that use more complex feature extraction methods such as the use of the brainmap and the empirical mode decomposition method. Such methods were able to reach better accuracy measures as they focused on analyzing the features of the dataset therefore, retrieved more relevant patterns associated with depression. Another disadvantage that can be pointed out regards the relatively limited dataset we used to train the model. In order to allow for more conclusive results, a larger dataset may need to test the model in more detail, focusing on its generalizability and robustness. Moreover, the research was restricted to the diagnosis of MDD only, and it is not clear if this approach would be appropriate for other mental disorders.

The findings from this study show that it is possible to train deep learning models for diagnosing major depressive disorder (MDD) using raw EEG signals and that it's not necessary to engineer rich features on these signals. This has the potential to reduce the development time of future systems for the detection of depression. However, these caveats which have been brought out in this study need further investigation together with the possible extension of this method to other psychiatric conditions.

5 Conclusion

This study aimed to develop a robust EEG-based classification model for Major Depressive Disorder (MDD) using Long Short-Term Memory (LSTM) networks relying solely on raw EEG signal data. The results demonstrated that the LSTM model achieved an accuracy of 76% in distinguishing between individuals with MDD and healthy controls. This finding suggests that LSTM networks can effectively learn discriminative patterns from raw EEG data, providing a promising approach for MDD classification. Future

research could explore several avenues to further improve the model's performance. These include refining the LSTM architecture, incorporating additional EEG features or preprocessing techniques, and investigating hybrid approaches that combine raw data learning with more traditional feature engineering methods. By addressing these areas, we may be able to develop even more accurate and reliable EEG-based diagnostic tools for MDD.

References

1. Li Z, Ruan M, Chen J, Fang Y (2021) Major depressive disorder: advances in neuroscience research and translational applications. Neurosci Bull 37:863–880. https://doi.org/10.1007/s12264-021-00638-3
2. Cai H et al (2021) Prevalence of suicidality in major depressive disorder: a systematic review and meta-analysis of comparative studies. Front Psych 12:690130. https://doi.org/10.3389/FPSYT.2021.690130/BIBTEX
3. Malgaroli M, Calderon A, Bonanno GA (2021) Networks of major depressive disorder: a systematic review. Clin Psychol Rev 85:102000. https://doi.org/10.1016/j.cpr.2021.102000
4. WHO. Depressive disorder (depression). https://www.who.int/news-room/fact-sheets/detail/depression
5. Anggraini MS, Hidayat A, Fernandez RL (2022) Analysis of the causes of mental health disorders in pregnant women in developing countries: a systematic review. Int J Adv Health Sci Technology 2:309–320. https://doi.org/10.35882/IJAHST.V2I5.158
6. Kalin NH (2021) Anxiety, depression, and suicide in youth. American J of Psychiatry 178:275–279. https://doi.org/10.1176/appi.ajp.2020.21020186
7. Twenge JM (2020) Why increases in adolescent depression may be linked to the technological environment. Curr Opin Psychol 32:89–94. https://doi.org/10.1016/j.copsyc.2019.06.036
8. Niyonsenga J, Mutabaruka J (2021) Factors of postpartum depression among teen mothers in Rwanda: a cross- sectional study. J Psychosom Obstet Gynecol 42:356–360. https://doi.org/10.1080/0167482X.2020.1735340
9. Association A.P (2000) Diagnostic and statistical manual of mental disorders. 5th Edition, American Psychiatric Association
10. Pangman VC, Sloan J, Guse L (2000) An examination of psychometric properties of the mini-mental state examination and the standardized mini-mental state examination: implications for clinical practice. Appl Nurs Res 13:209–213. https://doi.org/10.1053/APNR.2000.9231
11. Siuly S, Khare SK, Bajaj V, Wang H, Zhang Y (2020) A computerized method for automatic detection of schizophrenia using EEG signals. IEEE Trans Neural Syst Rehabil Engineering 28:2390–2400. https://doi.org/10.1109/TNSRE.2020.3022715
12. Baumgartner C, Koren JP (2018) Seizure detection using scalp-EEG. Epilepsia 59:14–22. https://doi.org/10.1111/epi.14052
13. Supriya S, Siuly S, Wang H, Zhang Y (2023) Epilepsy detection from EEG using complex network techniques: a review. IEEE Rev Biomed Eng 16:292–306. https://doi.org/10.1109/RBME.2021.3055956
14. Wijayanto I, Sulistyo T.M, Pratama YJN, Rahmaniar TD, Safitri AS, Sa'idah S, Hadiyoso S, Wibowo RA, Ismanto RAK, Putri AA, Khasanah AN, Diliana FH, Azzahra S, Gadama M, Utami AT (2024) Unlocking early detection and intervention potential: analyzing visual evoked potentials (VEPs) in Adolescents/Teenagers with narcotics abuse tendencies from the Tel Unisba Neuropsychology EEG Dataset (TUNDA). J Electron Electromed Eng Med Informatics 6:445–458 . https://doi.org/10.35882/JEEEMI.V6I4.476

15. Rizki MA (2024) Classification of brain tumor based on shape and texture features and machine learning. Indones J Electron Electromed Eng Med Informatics 6. https://doi.org/10.35882/27236G49
16. Liu W, Jia K, Wang Z, Ma Z (2022) A depression prediction algorithm based on spatiotemporal feature of EEG signal. Brain Sci 12:630. https://doi.org/10.3390/brainsci12050630
17. Khosla A, Khandnor P, Chand T (2022) Automated diagnosis of depression from EEG signals using traditional and deep learning approaches: a comparative analysis. Biocybern Biomed Eng 42:108–142. https://doi.org/10.1016/J.BBE.2021.12.005
18. Tajali A (2024) The Impactness of SMOTE as imbalance class handling for myocardial infarction complication classification using machine learning approach with data imputation and hyperparameter. Indones J Electron Electromed Eng Med Informatics 6. https://doi.org/10.35882/IJEEEMI.V6I4.13
19. Avots E, Jermakovs K, Bachmann M, Päeske L, Ozcinar C, Anbarjafari G (2022) Ensemble approach for detection of depression using EEG features. Entropy 24:211. https://doi.org/10.3390/e24020211
20. Chandrika GN, Mitra A, Satheeswaran V, Chowdhury R, Kumar P, Glory E (2024) Deep learning based classification of ECG signals using RNN and LSTM mechanism. J Electron Electromed Eng Med Informatics 6:332–342. https://doi.org/10.35882/JEEEMI.V6I4.496
21. Du X et al (2022) An efficient LSTM network for emotion recognition from multichannel EEG signals. IEEE Trans Affect Comput 13:1528–1540. https://doi.org/10.1109/TAFFC.2020.3013711
22. Fraschini M, Demuru M, Crobe A, Marrosu F, Stam CJ, Hillebrand A (2016) The effect of epoch length on estimated EEG functional connectivity and brain network organisation. J Neural Eng 13:36015. https://doi.org/10.1088/1741-2560/13/3/036015
23. Heckbert P (1998) Fourier transforms and the fast Fourier transform (FFT) algorithm. Comput Graph (ACM) 2
24. Cai H et al (2022) A multi-modal open dataset for mental-disorder analysis. Sci Data 9:178. https://doi.org/10.1038/s41597-022-01211-x
25. da Silva Lourenço C, Tjepkema-Cloostermans MC, van Putten MJAM (2021) Efficient use of clinical EEG data for deep learning in epilepsy. Clin Neurophysiol 132:1234–1240. https://doi.org/10.1016/j.clinph.2021.01.035
26. Akhter R, Beyette FR (2024) Detection of unfocused EEG epochs by the application of machine learning algorithm. Sensors 24:4829. https://doi.org/10.3390/s24154829
27. Schaworonkow N, Nikulin VV (2019) Spatial neuronal synchronization and the waveform of oscillations: implications for EEG and MEG. PLoS Comput Biol 15:e1007055. https://doi.org/10.1371/JOURNAL.PCBI.1007055
28. Koskela A, Jälkö J, Honkela A (2020) Computing tight differential privacy guarantees using FFT. In: Proceedings of the 23rd international conference on artificial intelligence and statistics, PMLR, pp 2560–2569
29. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9:1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
30. Raza MR, Hussain W, Merigó JM (2021) Cloud sentiment accuracy comparison using RNN, LSTM and GRU. In: 2021 Innovations in intelligent systems and applications conference (ASYU), pp 1–5. https://doi.org/10.1109/ASYU52992.2021.9599044
31. Noh S-H (2021) Analysis of gradient vanishing of RNNs and performance comparison. Information 12:442. https://doi.org/10.3390/info12110442
32. Wang Y (2017) A new concept using LSTM neural networks for dynamic system identification. In: 2017 American control conference (ACC), pp 5324–5329. https://doi.org/10.23919/ACC.2017.7963782
33. Vennerød CB, Kjærran A, Bugge ES (2021) Long Short-term Memory RNN. http://arxiv.org/abs/2105.06756. https://doi.org/10.48550/arXiv.2105.06756

34. Alhagry S, Fahmy AA, El-Khoribi RA (2017) Emotion recognition based on EEG using LSTM recurrent neural network. Int J Adv Comput Sci Applications 8. https://doi.org/10.14569/IJACSA.2017.081046
35. Yin Y, Zheng X, Hu B, Zhang Y, Cui X (2021) EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM. Appl Soft Comput 100:106954. https://doi.org/10.1016/J.ASOC.2020.106954
36. Gheorghe M, Mihalache S, Burileanu D (2023) Using deep neural networks for detecting depression from speech. In: European signal processing conference, pp 411–415. https://doi.org/10.23919/EUSIPCO58844.2023.10289973
37. Shen J et al (2023) Exploring the intrinsic features of EEG signals via empirical mode decomposition for depression recognition. IEEE Trans Neural Syst Rehabil Eng 31:356–365. https://doi.org/10.1109/TNSRE.2022.3221962

Static Finite Element Analysis in 3D-Printed Insole with Thermoplastic Polyurethane (TPU) 75A for Individuals with Diabetics

Raden Achmad Candra Putra[1], Tri Riana Lestari[1(✉)], Muhammad I'Tikap[1], Syifa Fauziah[1], and Subyakto[2]

[1] Department of Orthotics and Prosthetics, Poltekkes Kemenkes Jakarta I, South Jakarta, Indonesia
triarianalestari@yahoo.co.id
[2] Badan Riset dan Inovasi Nasional, Jakarta, Indonesia

Abstract. Individuals with diabetic neuropathic foot need custom-made insoles specifically intended to alleviate excessive pressure on the plantar area in order to avoid issues. The utilization of three-dimensional (3D) printing technology enables the production of more precise, personalized, and efficient. Therefore, this study aims to find out the geometric shape of the insole is designed to produce an insole that is both sufficiently flexible and durable to withstand the static load from the patient's body weight in static with uniform pressure. The process of creating 3D-printed insoles geometry involves some steps, including 3D scanning of the foot and 3D modeling. Then the 3D model insole was simulated statically uniform pressures using FEA software PrePoMax CalculiX version 2.1.4. The simulation resulted in a maximum of 35.29 mm deformation in model A with no arch support underneath the arch and 6.086 mm on the lateral trimline of the heel. Von Misses stress were 1.881 MPa and 1.237 MPa, respectively. Two geometry model insoles for diabetic were developed. The design with full support underneath the arch shows better resistance to deformation.

Keywords: 3D-printed insole · diabetics · finite element · static loading

1 Introduction

Individuals with diabetic neuropathic foot need custom-made insoles specifically intended to alleviate excessive pressure on the plantar area in order to avoid issues [1–3]. The utilization of three-dimensional (3D) printing technology enables the production of more precise, personalized [4, 5], and efficient custom insoles in comparison to classical techniques [6]. The next phases facilitate the manufacturing of insoles that are thoroughly matched to the anatomical shape of the foot and the specific demands of individuals with diabetes. The use of custom insoles has been proven in preventing foot ulcers [7]. Systematic review studies have shown that custom insoles can reduce pressure on the soles of the feet in patients with diabetic peripheral neuropathy [1, 2, 8, 9] . Therefore, the use of footwear with custom insoles has become the standard of care for individuals with diabetes in Australia [10].

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 485–493, 2026.
https://doi.org/10.1007/978-981-95-9233-3_35

Along with the development of science and technology, the use of design and manufacturing techniques with 3D designing and printing processes is being developed in medical science, including orthotics prosthetics [11]. The orthotist can directly design the shape of the insole according to the patient's foot by considering the contour of the scan result and the location point of the excess pressure through software, then print it using a 3D printer on the same day. Therefore, the process of making insoles can be faster [12]. In addition, reducing the use of gypsum waste and plastics that are difficult to decompose can reduce the adverse impact on the environment and the efficiency of the materials used makes the price of custom insoles with this 3D printing method more affordable. Therefore, the purpose of this study is to develop customized insoles with 3d printing technology to prevent foot ulcers and amputations in individuals with diabetes.

The technique of static finite element analysis (FEA) is used to analyze the distribution of stress in 3D-printed insoles specifically developed for diabetes patients [13]. This method enables the simulation of the interactions between various forces and the insole structure [14]. By fine-tuning insole design using Finite Element Analysis (FEA), it is feasible to guarantee efficient pressure relief at crucial regions, hence greatly decreasing the likelihood of ulcers, a prevalent issue for persons with diabetes. The selection of material is important in this optimization process. Material of 75A thermoplastic polyurethane (TPU) is suggested due to its ideal combination of rigidity and flexibility, which improves both the performance and comfort of the insoles during finite element analysis (FEA) evaluations [15]. Recent research emphasizes that the latest developments in 3D-printing and material science show potential for enhancing the offloading capacities of these insoles, as well as customizing material characteristics to meet individual patient requirements [16].

The characteristics of 75A TPU, which is flexible and rigid, may improve comfort during prolonged wear making it an ideal material for insoles designed for diabetic patients as its biocompatibility [17]. The durability and resilience of 75A TPU ensure that it can withstand the demands of daily activities while providing adequate support and shock absorption. Those characteristics are important for diabetic patients who require reliable cushioning to protect their feet from injury. The lightweight nature of 75A TPU also contributes to ease of movement, promoting overall foot health by enabling patients to be able to do regular activities without feeling uncomfortable.

This study aims to find out the geometric shape of the insole is designed to produce an insole that is both sufficiently flexible and durable to withstand the static load from the patient's body weight in static with uniform pressure with 75A TPU. This study will contribute to the development of insole's design using the digital technology and give further insight on the ideal shape of the custom-made insole for optimal durability based on uniform pressure through simulation. Additionally, the findings may have implications for the development of a more effective insole in redistributing plantar pressure in diabetic patient.

2 Methods

The process of creating 3D-printed insoles geometry involves some steps, including 3D scanning of the foot and 3D modeling. These steps enable the production of insoles that can be customized to the shape of the foot. Digital foot scanning aims to obtain a precise

3D mesh image of the patient's foot for which the insole will be created as shown in Fig. 1. In the production of 3D-printed insoles, scanning is performed using a handheld 3D scanner that employs hybrid infrared and ultra-bright LED white structured light technology [18].

(a) (b)

Fig. 1. Scanning process. (a) Handheld EinScan H scanner, (b) Process of foot scanning with handheld 3D scanner in non-weight bearing

The 3D mesh image of the foot obtained from the 3D scanner is then used to create a 3D model. The insole design is developed according to the patient's needs for pressure distribution using Mesh mixer software. This allows the insole to be designed with precision and exported as an STL file, which is then ready for the simulation process. The two designed of insole were simulated using FEA software PrePoMax CalculiX version 2.1.4. The initial simulation for the insole designed would be in the form of static uniform pressure. The calculation of magnitude uniform pressure for the simulation was calculated using pressure as Eq. (1);

$$\mathrm{P} = \frac{\frac{Patient's weight}{2}\ safety\ factor.g}{A} \tag{1}$$

where the weight of patient is 60 kg, where the load of an individual foot will be divided by two, with 1.2 safety factor, and 0.013 m^2 surface area, which is calculated as Eq. (2):

$$\mathrm{P} = \frac{\frac{60\ kg}{2}\ (1.2)\ \left(9.81\ m/s^2\right)}{0.013\ m^2} = 27\ kPa \tag{2}$$

The simulation uses thermoplastic polyurethane (TPU) 75 A as the insole material. TPU has been becoming a common material used in 3D printing with fused deposition

modeling (FDM) [19] . The material provides a balance of flexibility and durability [20]. Therefore, it is ideal as insole material [21]. Its ability to evenly distribute pressure can help to minimize discomfort and prevent further foot complications. The properties of TPU material 75 A are 11 MPa Young's modulus, 0.45 Poisson's ratio, and 22 MPa tensile strength. These properties make TPU 75 A suitable choice for insole production as it can provide the necessary support and cushioning for the patient.

3 Result

There are two geometry models developed. One model (Model A) without support in the arch area of the foot, and one model (Model B) with full support in the arch area shown in Fig. 2. All models were developed with three quarter length to accommodate for push-off.

(a) Model A (b) Model B

Fig. 2. Two 3d designed model's insole. (a). Insole design without support in the arch, (b). Insole design with full support in the arch

The solid mesh model of the insole has 952 elements and 2047 nodes. The simulation mimics how the insole will behave when it bears a static loading during standing. Then, a floor model is also incorporated within the simulation with a 2100 GPa Young's modulus. The floor is conditioned to be fixed in all degrees of freedom and a uniform pressure from a foot is applied on the top surface of the insole (blue coloured area in Fig. 3. There is also a contact area between the bottom surface of the insole and the surface of the floor. The magnitude of the uniform pressure is 27 kPa. The magnitude is calculated from a hypothetical 60 kg patient where the load of an individual foot will be divided by 2, with 1.2 safety factor, and 0.013 m^2 surface area is measured from the mesh model approximation in PrePoMax.

The simulation resulted in a maximum of 35.29 mm in model A at arch of insole and 6.08 mm in model B deformation as shown in Fig. 4. Von Mises stress results shown in Fig. 5 indicated that model A experienced higher stress levels 1.881 MPa as compared to model B 1.237 MPa.

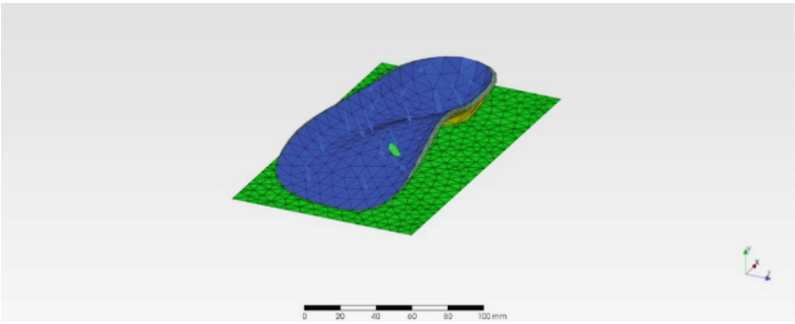
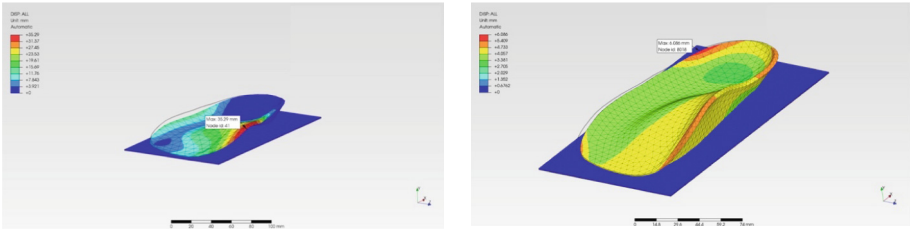
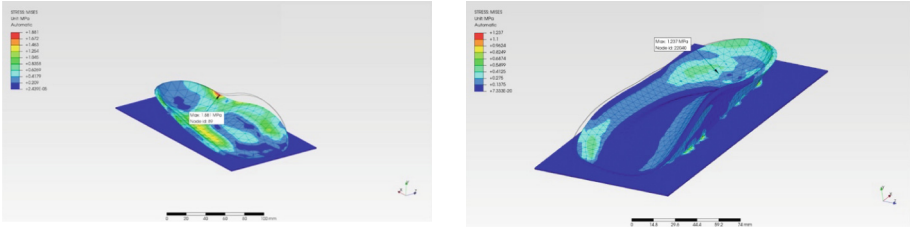


Fig. 3. The boundary conditions of the FEA simulation



(a) (b)

Fig. 4. Deformation map of result of the simulation (a) models a with no support under the arch, and (b) model b with full support under arch



(a) (b)

Fig. 5. Stress map result of the simulation (a) front view, (b) side view

4 Discussion

Two geometry model insoles for diabetics were developed. The FEA simulation of the insole model A results in a 35 mm maximum strain from its initial position in the arch area. However, model B shows better results in maximum strain with only 6.086 mm in the lateral trimline of the heel. However, in this simulation, the simulated uniform pressure is not the most accurate representation of real plantar pressure since

non-simulated plantar pressure on a flat surface tends to be spreader and focused on several concentrated areas. The maximum stress from the simulated pressure was only 1.881 MPa in model A and 1.237 MPa in model B. This is still less than the material's tensile strength of 22 MPa, which means that the deformation did not cause the material to break. These results show that the materials used in both models A and B are suitable for diabetic patients in terms of withstanding pressure and deformation. However, model B shows better resistance to deformation than model A.

There are other studies that evaluate the design and material of 3D-printed insoles for reducing peak plantar pressure for diabetic patients [22, 23]. The materials used in making 3D-printed insoles are varied and include thermoplastic elastomers [22, 24], ABS composite combined with silicone [25, 26] and polyurethane [15, 27–29] . Despite all different material uses, 3D-printed insoles show potential benefits of customized orthotic devices in preventing diabetic foot complications.

Our novel 3D-printed insole design with arch support could provide total contact on the plantar surface, which could redistribute plantar pressure evenly and thus prevent ulcers. A study conducted by Niu et al. [30] shows that the insole with arch support design could reduce the pressure by 48 kPa by finite element analysis. Another study by Zuniga et al. [31] found that 3D-printed insoles with two flexible polymers with hard shore A 70 and 85 could have the same effectiveness in reducing peak plantar pressure as the conventional method with handmade. However, the time required for production and cost are significantly lower in the 3D printed method.

Even though the potential benefits of 3D printing in terms of design and material use, there were still some weaknesses in our study. One of the major limitations was that the simulation conducted in this study was with only static uniform pressure, which did not accurately represent the dynamic and varying pressures that occurred during actual walking. Additionally, this study was only simulated using one type of commonly available filament material in Indonesia, which may limit the generalizability of the results to other materials.

Our study emphasizes how 3D-printed insoles can play a positive role in caring for diabetic feet. Model B's enhanced resistance to deformation indicates that with thoughtful design, insoles can provide exceptional comfort and help distribute pressure evenly, which is crucial for preventing foot ulcers in individuals with diabetes. The ability to customize insoles with 3D printing not only enhances their effectiveness but also makes them easier to manage and quicker to produce than conventional methods. Even though this study has shown potential results, it is important for further research to look into potential 3D printing materials [32] and test them in real walking situations to truly understand how we can improve these insoles for diabetic foot health.

5 Conclusion

The aim of this study to find out the geometric shape of the insole is designed to produce an insole that is both sufficiently flexible and durable to withstand the static load from the patient's body weight in static with uniformed pressure with 75A TPU. Two geometry model insoles for diabetic were developed. The design could be made in more precise in redistributing plantar pressure and offloading specific area for patient with ulcer.

The simulation of the insole model B with full support underneath the arch shows better results in maximum strain with only 6.086 mm in the lateral trimline of the heel and the maximum stress was 1.237 MPa. The future work should be executing the simulation with other common and available filaments materials such as PLA, PETG, etc. Another loading conditions to be simulated not only static loading during standing but also dynamic loading during walking, specifically at the initial contact phase, loading response phase, and mid-stance phase of gait.

References

1. Paton J, Bruce G, Jones R, Stenhouse E (2011) Effectiveness of insoles used for the prevention of ulceration in the neuropathic diabetic foot: a systematic review. https://doi.org/10.1016/j.jdiacomp.2009.09.002
2. Collings R, Freeman J, Latour JM, Paton J (2021) Footwear and insole design features for offloading the diabetic at risk foot—A systematic review and meta-analyses. https://doi.org/10.1002/edm2.132
3. Paton J, Glasser S, Collings R, Marsden J (2016) Getting the right balance: insole design alters the static balance of people with diabetes and neuropathy. J Foot Ankle Res 9. https://doi.org/10.1186/s13047-016-0172-3
4. Völz D, Schneider J, Wuttke U (2023) Digital personalized medical insole process. Curr Dir Biomed Eng 9. https://doi.org/10.1515/cdbme-2023-1008
5. Völz D, Schneider J, Wuttke U (2023) Digital personalized medical insole process a step forward to more personalized medical support. Curr Dir Biomed Eng. https://doi.org/10.1515/cdbme-2023-1008
6. Mamo HB, Adamiak M, Kunwar A (2023) 3D printed biomedical devices and their applications: a review on state-of-the-art technologies, existing challenges, and future perspectives. https://doi.org/10.1016/j.jmbbm.2023.105930
7. Jandova S, Mendricky R (2022) Benefits of 3D printed and customized anatomical footwear insoles for plantar pressure distribution. 3D Print Addit Manuf 9. https://doi.org/10.1089/3dp.2021.0002
8. Korada H, Maiya A, Rao SK, Hande M (2020) Effectiveness of customized insoles on maximum plantar pressure in diabetic foot syndrome: a systematic review. https://doi.org/10.1016/j.dsx.2020.06.041.
9. Joeliantina A, Hamid IN, Sulistyowati E (2023) The experiences of patients with diabetes and strategies for their management during the COVID-19 pandemic: a qualitative study. Int J Adv Health Sci Technol 3. https://doi.org/10.35882/IJAHST.V3I1.214
10. van Netten JJ et al (2018) Diabetic foot Australia guideline on footwear for people with diabetes. J Foot Ankle Res. 11. https://doi.org/10.1186/s13047-017-0244-z
11. Choo YJ, Boudier-Revéret M, Chang MC (2020) 3D printing technology applied to orthosis manufacturing: narrative review. https://doi.org/10.21037/apm-20-1185
12. Davia-Aracil M, Hinojo-Pérez JJ, Jimeno-Morenilla A, Mora-Mora H (2018) 3D printing of functional anatomical insoles. Comput Ind 95. https://doi.org/10.1016/j.compind.2017.12.001
13. Peker A et al (2020) Additive manufacturing and biomechanical validation of a patient-specific diabetic insole. Polym Adv Technol 31. https://doi.org/10.1002/pat.4832
14. Ragab A (2023) Modeling orthotic insole for flatfoot using finite element analysis and 3D scanning. Presented at the. https://doi.org/10.46254/eu05.20220261

15. Iacob MC, Popescu D, Petcu D, Marinescu R (2023) Assessment of the flexural fatigue performance of 3D-printed foot orthoses made from different thermoplastic polyurethanes. Appl Sci (Switz) 13. https://doi.org/10.3390/app132212149
16. Shaikh S, Jamdade B, Chanda A (2023) Effects of customized 3D-printed insoles in patients with foot-related musculoskeletal ailments—a survey-based study. PRO 5. https://doi.org/10.3390/prosthesis5020038
17. Haryńska A, Gubanska I, Kucinska-Lipka J, Janik H (2018) Fabrication and characterization of flexible medical-grade TPU filament for fused deposition modeling 3DP technology. Polymers (Basel) 10. https://doi.org/10.3390/polym10121304
18. Najafabadi DF, Rezaie MR, Forghany S (2020) The validity and reliability of a low-cost handheld 3D scanner for use in orthotics and prosthetics. J Rehabil Sci Res 7. https://doi.org/10.30476/jrsr.2019.82703.1034
19. Shin EJ, Jung YS, Park CH, Lee S (2023) Eco-friendly TPU/PLA blends for application as shape-memory 3D printing filaments. J Polym Environ 31. https://doi.org/10.1007/s10924-023-02799-w
20. Shin EJ, Park Y, Jung YS, Choi HY, Lee S (2022) Fabrication and characteristics of flexible thermoplastic polyurethane filament for fused deposition modeling three-dimensional printing. Polym Eng Sci 62. https://doi.org/10.1002/pen.26075
21. Tao Y, Shao J, Li P, Shi SQ (2019) Application of a thermoplastic polyurethane/polylactic acid composite filament for 3D-printed personalized orthosis. Materiali in Tehnologije 53. https://doi.org/10.17222/MIT.2018.180
22. Muir BC, Li JS, Hudak YF, Kaufman GE, Cullum S, Aubin PM (2022) Evaluation of novel plantar pressure-based 3-dimensional printed accommodative insoles—a feasibility study. Clin Biomech 98. https://doi.org/10.1016/j.clinbiomech.2022.105739
23. Mancuso M, Bulzomì R, Mannisi M, Martelli F, Giacomozzi C (2023) 3D-printed insoles for people with type 2 diabetes: an Italian, ambulatory case report on the innovative care model. Diabetology 4. https://doi.org/10.3390/diabetology4030029
24. Hudak YF et al (2022) A novel workflow to a fabricate patient-specific 3D printed accommodative foot orthosis with personalized latticed metamaterial. Med Eng Phys 104. https://doi.org/10.1016/j.medengphy.2022.103802
25. Chhikara K, Gupta S, Chanda A (2022) Development of a novel foot orthosis for plantar pain reduction. In: Materials today: proceedings. https://doi.org/10.1016/j.matpr.2022.04.361
26. Putra KB, Wood H, Wensman J, Leonard JA, Shih A (2023) A custom silicone foot orthosis for partial Calcanectomy fabricated using a 3D-printed Mold. J Prosthet Orthot 35. https://doi.org/10.1097/JPO.0000000000000441
27. Iacob MC, Popescu D, Baciu F (2023) Effect of process parameters on the hardness of 3D-printed thermoplastic polyurethane that includes foaming agent. Materiale Plastice 60. https://doi.org/10.37358/MP.23.4.5694
28. Kumar R, Sarangi SK (2023) On the enhanced mechanical characteristics of 3D printed architected spiderweb lattice structures through overall density variation. Eng Res Express 5. https://doi.org/10.1088/2631-8695/acf981
29. Orsu B, Shaik YP (2022) Compression strength analysis of customized shoe insole with different infill patterns using 3D printing. OAlib 09. https://doi.org/10.4236/oalib.1108712
30. Niu J, Liu J, Zheng Y, Ran L, Chang Z (2020) Are arch-conforming insoles a good fit for diabetic foot? Insole customized design by using finite element analysis. Hum Factors Ergon Manuf 30. https://doi.org/10.1002/hfm.20841

31. Zuñiga J et al (2022) Development of 3D-printed Orthopedic insoles for patients with diabetes and evaluation with electronic pressure sensors. Designs (Basel) 6. https://doi.org/10.3390/designs6050095
32. Kristiawan RB, Imaduddin F, Ariawan D, Ubaidillah Arifin Z (2021) A review on the fused deposition modeling (FDM) 3D printing: filament processing, materials, and printing parameters. https://doi.org/10.1515/eng-2021-0063

Ten Hand Motion Classification Using Autoregression Feature Extraction Based EMG Signal

Triwiyanto[1(✉)], Endro Yulianto[1], Anita Miftahul Maghfiroh[1], and Wahyu Caesarendra[2]

[1] Department of Medical Electronics Technology, Poltekkes Kemenkes Surabaya, Surabaya, Indonesia
triwiyanto@ieee.com

[2] Faculty of Mechanical Engineering, Opole Unviersity of Technology, Opole, Poland

Abstract. The primary problem addressed in this study is the challenge of accurately classifying hand motions using Electromyography (EMG) signals, which are inherently complex due to the overlapping electrical activities of flexor and extensor muscles. The aim of the research is to develop a robust classification model for ten distinct hand gestures by employing autoregressive feature extraction in combination with various classifiers, including K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Linear Discriminant Analysis (LDA). The methodology involved recording EMG signals from two channels and applying autoregressive modeling to extract relevant features. The classifiers were then evaluated for their performance based on classification accuracy. The results revealed that LDA achieved the highest accuracy of 80% at an autoregressive order of 12, while KNN's performance declined to an accuracy of 50%, indicating its sensitivity to feature complexity. Additionally, the confusion matrix showed that while certain gestures, such as "CLO," were classified perfectly, others like "IND" and "THI" experienced misclassification rates of 30% and 46%, respectively. In conclusion, while LDA proved effective in classifying EMG signals for hand gestures, the variability in accuracy highlights the need for further refinement in feature extraction and potentially the exploration of more advanced classifiers to improve recognition capabilities in practical applications such as prosthetics and rehabilitation engineering.

Keywords: hand motion · autoregression · SVM · LDA · KNN

1 Introduction

Electromyography (EMG) signals have become a critical tool in the classification of hand motions, especially for applications such as prosthetics control, rehabilitation devices, and human-computer interaction [1–4]. Despite recent advances, achieving accurate and reliable hand motion recognition remains a challenge due to the complexity of muscular signals and the non-linear nature of EMG data [5–8]. The challenge is further compounded by issues such as signal noise, sensor placement variability, and individual

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 494–506, 2026.
https://doi.org/10.1007/978-981-95-9233-3_36

differences in muscle anatomy [9–11]. In light of this, there is a pressing need for robust classification methods that can reliably interpret subtle EMG patterns. The dataset utilized in this research, consisting of multiple hand movements like pronation, supination, and wrist flexion, offers a rich set of signals that can contribute to more accurate classification. However, to address these challenges, the development of effective feature extraction techniques is paramount [9–11].

State-of-the-art methods for EMG-based motion classification largely rely on machine learning techniques combined with advanced feature extraction methods. Techniques such as Convolutional Neural Networks (CNNs) [12, 13], Support Vector Machines (SVM) [14], and Long Short-Term Memory (LSTM) [15, 16] networks have demonstrated remarkable performance in classifying complex motion patterns. Autoregression (AR) modeling, a statistical technique that captures temporal dependencies in time-series data, has also been extensively applied to EMG signals for feature extraction [17]. AR models excel in transforming raw EMG signals into meaningful features that reduce noise while enhancing the essential patterns needed for classification [17–19]. However, while these methods have proven effective, there remains a need to refine their precision and adaptability, especially when dealing with multi-class problems involving subtle hand motions.

A key research gap exists in the classification of multiple hand motions using EMG signals, particularly when leveraging autoregressive models as the primary feature extraction method. While existing studies have demonstrated the effectiveness of AR modeling in extracting features, few have focused on its application in the classification of a wide array of hand motions simultaneously [17, 18, 20]. Additionally, much of the current literature focuses on the classification of simpler motions or binary tasks, leaving the classification of more complex multi-class problems, such as ten different hand motions, relatively underexplored. Moreover, the integration of autoregression with advanced machine learning models like deep neural networks for such tasks is still evolving, creating a gap that this research aims to fill.

The primary aim of this research is to develop a robust classification model capable of distinguishing ten hand motions based on EMG signals by utilizing autoregression-based feature extraction. By applying autoregressive modeling to a rich dataset of hand motions, this study seeks to improve the accuracy of motion classification and reduce the computational complexity typically associated with high-dimensional EMG data. The proposed model will be evaluated for its ability to generalize across multiple participants and sensor placements, with the ultimate goal of improving the precision and reliability of hand motion classification.

The main contribution of this research lies in the integration of autoregressive feature extraction with state-of-the-art machine learning techniques to enhance the classification of multiple hand motions. By leveraging a comprehensive dataset of EMG signals and focusing on ten specific hand movements, this work will contribute to the growing field of EMG-based motion recognition. Additionally, the findings may have broader implications for the design of more responsive prosthetic devices, improved rehabilitation tools, and innovative human-computer interaction systems. In summary, this research fills an existing gap by addressing multi-class hand motion classification with a novel combination of feature extraction and classification techniques.

2 Method

2.1 Dataset

The dataset used in this study was obtained from a publicly available biosignals repository. It consists of EMG data collected from eight participants: six males and two females, aged between 20 and 35 years, all of whom were normally limbed without any neurological or muscular disorders. Each subject was asked to perform a series of finger movements, following a predefined protocol, to generate EMG signals. The recruited subjects provided informed consent in accordance with ethical guidelines before participation. This dataset was chosen due to its comprehensiveness in capturing finger movements across a diverse group of subjects, ensuring variability in signal patterns while maintaining controlled experimental conditions. The focus of the data collection was on extracting EMG signals relevant to various hand and finger movements, providing a robust foundation for further analysis.

2.2 Data Collection

The data collection process involved setting up subjects in a controlled environment where their arm position was stabilized to reduce the effect of limb movement on the recorded EMG signals. Subjects were seated on an armchair, and their arms were fixed in position to minimize interference from posture changes. The EMG signals were recorded using two high-precision Delsys DE 2.x series EMG sensors. These sensors were chosen for their sensitivity and ability to capture minute electrical activity in the muscles, which is crucial for accurate motion classification. To ensure stable contact and reduce motion artifacts, a 2-slot adhesive skin interface was applied to each sensor, securing them to the skin. The data was then processed through the Bagnoli Desktop EMG System from Delsys Inc., a widely recognized system in EMG research for its accuracy and high-quality signal processing capabilities. Each subject performed a series of predefined hand and finger movements, generating EMG signals for further analysis.

2.3 Data Processing

Once the EMG signals were collected, they were processed to remove noise and artifacts [21]. Raw EMG data is often noisy due to factors such as muscle fatigue, motion artifacts, and external interference. To address this, a band-pass filter was applied to the EMG signals to retain frequencies between 20 Hz and 450 Hz [22–24], which is the range most relevant to muscle activity, while filtering out unwanted noise and high-frequency artifacts. Next, the data was normalized to account for variability between subjects, such as differences in muscle size, strength, and sensor placement. Autoregressive (AR) modeling was employed as the primary feature extraction method. The AR model is well-suited for time-series data like EMG because it captures the temporal dependencies between signal points. The AR coefficients were computed for each signal window, generating feature vectors for each movement class that could be used for classification.

2.4 Data Analysis

The extracted AR features were used to classify the different hand and finger movements. To analyze the data, a machine learning model was developed, focusing on supervised learning algorithms such as Support Vector Machines (SVM) and neural networks, which have been proven effective in classifying EMG data. The dataset was split into training and testing sets, with the training set used to train the classification model and the testing set to evaluate its performance. Cross-validation was employed to ensure the robustness of the model and to prevent overfitting. Performance metrics such as accuracy, precision, recall, and F1-score were used to evaluate the classifier's effectiveness. Special attention was given to the multi-class nature of the problem, as the dataset involved ten distinct hand motions, which added complexity to the classification task. The model's performance was compared with state-of-the-art methods to determine its relative efficacy in classifying hand and finger movements based on EMG signals.

Electromyography (EMG) signals represent electrical activity in muscles, often used in medical and rehabilitation applications. Feature extraction from EMG signals is essential for analyzing muscle function and controlling prosthetic devices. Autoregression (AR) is a powerful method for this, where the current value of a signal is expressed as a linear function of previous values. AR modeling simplifies EMG signal representation, capturing underlying patterns and reducing noise. For an AR model of order p, the signal x(n) is described as Eq. (1):

$$\mathrm{x(n)} = \sum_{i=1}^{p} a_i x(n-i) + e(n) \tag{1}$$

where a_i are the AR coefficients and e(n) is the prediction error. Autorepression as a feature extraction technique in electromyography (EMG) signal processing works by modifying the original signal in a controlled feedback loop to reduce noise and highlight relevant data. This method emphasizes the most significant parts of the EMG signal by applying a repression function, typically with a proportional constant, to the signal. The aim is to eliminate redundancy and irrelevant features, ensuring that only meaningful information from muscle activity is retained. This process is essential in applications like bionic prostheses or neuromuscular diagnostics, where clear signal interpretation is critical. Autorepression is often combined with other methods, such as wavelet-based techniques like the Discrete Wavelet Transform (DWT), to further enhance precision in identifying muscle activation patterns, thus offering a robust approach to handling the complexity of EMG data.

3 Result

Figure 1 represents the Electromyography (EMG) signals captured from two channels, which correspond to the flexor and extensor muscles in the forearm. EMG signals are recorded during ten distinct hand motions, each motion generating a unique pattern of electrical activity based on muscle contraction and relaxation. The data from the two channels—flexor and extensor—show complementary behaviors since these muscles work in opposition to control hand movement. In the figure, Channel 1 (depicted in blue) corresponds to the flexor muscles, and Channel 2 (depicted in red) represents the

extensor muscles. The x-axis shows the time in seconds, indicating the duration over which the signals were recorded, while the y-axis displays the amplitude of the EMG signal, reflecting the intensity of muscle activity in microvolts (μV). The signal patterns reflect various hand motions, including flexion and extension of different fingers and the wrist, such as radial and ulnar deviation, pronation, and supination. The fluctuations in the signal amplitude indicate muscle activation corresponding to the movements. Larger peaks suggest higher muscular contraction, typical of more forceful motions, while the smaller variations are indicative of rest or subtle muscle activity. The two signals overlap in some instances, which demonstrates simultaneous contraction of both flexor and extensor muscles during certain complex motions. This data is crucial for classifying hand movements in applications such as prosthetics control, human-computer interaction, and rehabilitation engineering.

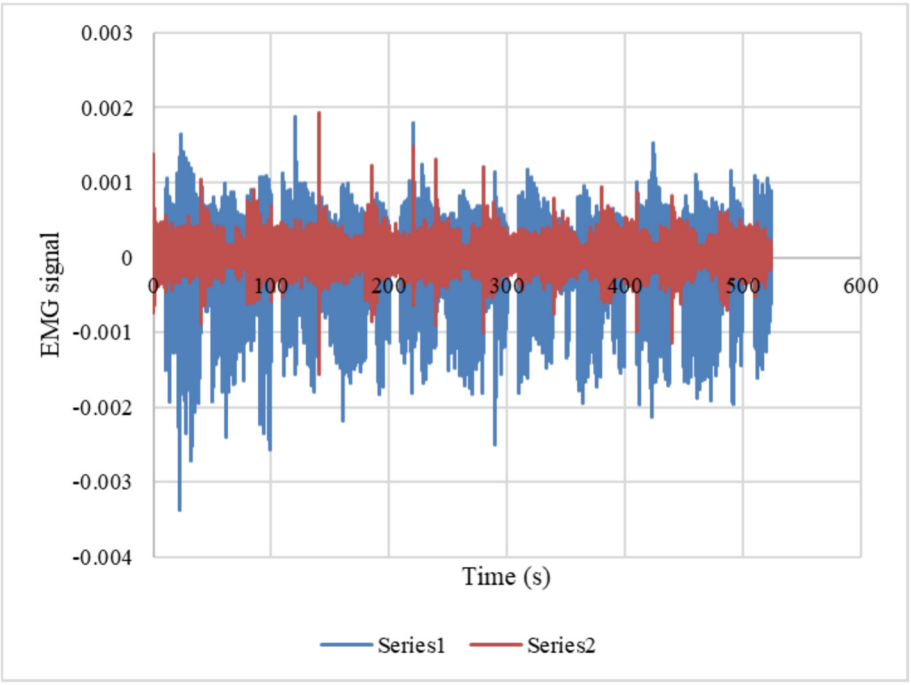


Fig. 1. EMG signal representation from ten motion from two channel EMG (Fig. 2)

Figure 2 illustrates the accuracy performance of different machine learning classifiers, including K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Linear Discriminant Analysis (LDA), across varying autoregressive (AR) model orders (2, 4, 6, 8, 10, and 12). On the horizontal axis, the classifiers are grouped under different AR orders, while the vertical axis represents the classification accuracy ranging from 0 to 0.9. Each classifier's performance changes based on the AR order used in feature extraction from electromyography (EMG) signals. At AR order 2, all three classifiers show moderate performance, with LDA slightly outperforming the other two classifiers,

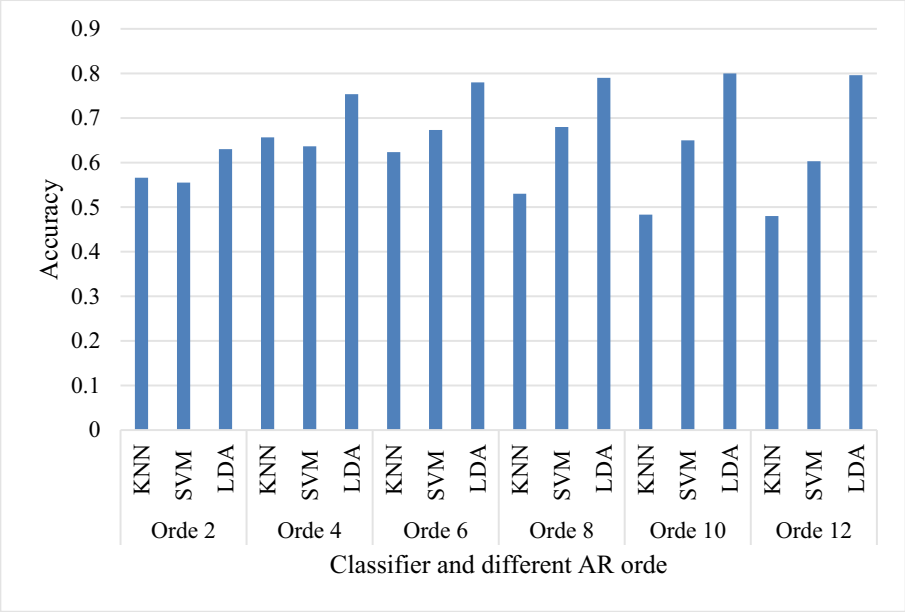


Fig. 2. Accuracy of three classifier (KNN, SVM, LDA) with auto regression feature extraction for orde (2, 4, 6, 8, 10, 12)

achieving an accuracy around 0.65, while KNN shows the lowest accuracy at approximately 0.55. This suggests that even at lower AR orders, LDA is able to extract useful features from the data for classification. When the AR order is increased to 4, KNN's accuracy improves considerably, reaching approximately 0.7, slightly outperforming both SVM and LDA, which also see a marginal improvement in accuracy. This trend continues as we move to AR order 6, with LDA achieving its highest accuracy so far, close to 0.75, while KNN and SVM remain consistent, indicating that the increase in AR order helps improve feature extraction for most classifiers. However, a shift occurs at AR order 8, where KNN experiences a notable decrease in accuracy, falling to around 0.6, while SVM and LDA maintain their performance. This suggests that increasing the AR order beyond a certain point may introduce redundancy or noise that negatively impacts the performance of certain classifiers, such as KNN. At AR order 10, the performance of all classifiers decreases, with KNN dropping further to approximately 0.5. Both SVM and LDA also experience slight declines, indicating that AR order 10 might introduce less relevant features or too much complexity for the classifiers to handle effectively. Finally, at AR order 12, LDA achieves the highest accuracy of any classifier across all AR orders, reaching around 0.8. This shows that LDA is particularly effective at handling higher-order AR features, while KNN and SVM struggle at this level, with their accuracy stagnating or slightly improving compared to the previous AR order. In conclusion, Fig. 3 highlights that while increasing AR order improves accuracy up to a certain point, it can also lead to a drop in performance for certain classifiers like KNN, while LDA continues to show strong performance, especially at higher AR orders.

	CLO	IND	LIT	MID	RIN	THI	THL	THM	THR	THU	
CLO	1	0	0	0	0	0	0	0	0	0	
IND	0	0.7	0	0.1	0	0.1	0	0	0	0.1	1.0
LIT	0	0	1	0.1	0	0.1	0.1	0.06	0	0	
MID	0	0.1	0	0.8	0	0	0	0.03	0	0	
RIN	0	0	0	0	0.9	0	0	0	0.1	0	
THI	0	0.1	0	0.1	0.1	0.7	0	0.15	0	0	
THL	0	0	0	0	0	0	0.8	0.14	0	0	
THM	0	0	0	0	0	0.2	0	0.72	0	0	0.0
THR	0	0	0	0	0.2	0	0	0	0.8	0	
THU	0	0.1	0	0	0	0	0	0	0	0.8	

Fig. 3. Confusion matrice for 10 class based on LDA classifier with auto regression feature orde 10

The confusion matrix presented represents the performance of a Linear Discriminant Analysis (LDA) classifier applied to a ten-class classification problem using autoregression features with an order of 10 (Fig. 3). In this context, the confusion matrix is a valuable tool for visualizing how well the classifier distinguishes between the ten different classes, which are labeled as "CLO," "IND," "LIT," "MID," "RIN," "THI," "THL," "THM," "THR," and "THU." Each of these labels likely corresponds to distinct hand gestures or finger motions, as is common in electromyographic (EMG) signal classification tasks. The rows of the matrix represent the true classes (actual gestures), while the columns represent the predicted classes. The values within the matrix indicate the proportion of instances that were correctly or incorrectly classified by the LDA model.

The diagonal elements in the confusion matrix (from top-left to bottom-right) show the correct classifications, i.e., where the predicted class matches the true class. For example, the "CLO" class was classified with an accuracy of 1.0, indicating perfect performance for that gesture. Similarly, the "LIT," "MID," "RIN," "THL," "THM," "THR," and "THU" gestures were also classified with high accuracies, ranging from 0.72 to 0.9. These high diagonal values reflect that the classifier performed well on these classes, meaning it could distinguish these gestures effectively from others.

However, there are off-diagonal elements that represent misclassifications, where the model predicted a different class than the actual one. For instance, the "IND" class shows a 0.7 value on the diagonal, meaning that 70% of the "IND" gestures were correctly classified. However, there are misclassifications: 10% of the "IND" gestures were classified as "MID," and smaller percentages were misclassified as "THI" and "THU." These off-diagonal values reflect the challenges the classifier faced in differentiating certain classes, which could be due to overlapping features or similarities in the gestures themselves.

Another key observation is that certain gestures seem to have been confused more frequently than others. For example, "THI" was sometimes misclassified as "THL" or "THM," and vice versa. This could indicate that these gestures have similar feature

representations, making it harder for the LDA classifier to distinguish between them. It's worth noting that the gestures "CLO" and "LIT" were classified with perfect accuracy, indicating that these gestures are likely distinct enough in terms of their autoregression features for the classifier to easily recognize them.

The overall structure of the confusion matrix suggests that while the LDA classifier with autoregression features is effective at distinguishing many of the gestures, there is still room for improvement in certain areas, particularly for classes like "IND," "THI," and "THM," where misclassification rates are relatively higher. These errors could be addressed by refining the feature extraction process, such as by exploring different orders of autoregression or incorporating additional signal features beyond autoregression coefficients. Alternatively, employing more complex machine learning classifiers might improve performance for these ambiguous cases, as LDA is a linear classifier and may struggle with more complex, non-linear decision boundaries.

In conclusion, the confusion matrix offers valuable insights into the classifier's performance, highlighting both its strengths and areas for improvement. While the LDA classifier with autoregression features of order 10 performs well for several classes, further investigation is needed to address misclassifications in specific gestures. This could involve tuning the classifier, exploring alternative feature sets, or even testing more sophisticated classification algorithms to enhance the accuracy and robustness of the hand gesture recognition system.

Table 1. Classification report for LDA classifier with AR orde 10

	Report			
Classes	precision	recall	f1-score	support
CLO	1	1	1	21
IND	0.75	0.72	0.74	29
LIT	0.79	0.64	0.71	36
MID	0.79	0.87	0.83	31
RIN	0.82	0.94	0.87	33
THI	0.54	0.65	0.59	20
THL	0.86	0.83	0.84	29
THM	0.7	0.72	0.71	32
THR	0.93	0.83	0.88	30
THU	0.82	0.79	0.81	39

The classification performance of a model designed for hand motion recognition using Electromyography (EMG) signals is measured through precision, recall, F1-score, and support (Table 1). Each metric provides insights into the model's effectiveness in distinguishing among different classes. For instance, the class CLO achieved perfect metrics, indicating that it correctly identified all instances with no errors. In contrast, other classes like IND and LIT showed moderate performance, with F1-scores of 0.74 and

0.71 respectively, suggesting that while they performed reasonably well, there is room for improvement. The MID and RIN classes exhibited stronger metrics, with RIN showing a high recall of 0.94, demonstrating its ability to identify most actual instances. However, the THI class highlighted a significant issue, with lower precision and recall, reflecting challenges in correctly classifying instances of this class. The F1-scores across the classes range from 0.59 for THI to 1.00 for CLO, indicating varied performance levels. Overall, while some classes performed excellently, others revealed areas needing enhancement, particularly through potential model tuning and data augmentation strategies to improve classification accuracy.

4 Discussion

The results of this study provide significant insights into the classification of hand motions using Electromyography (EMG) signals, particularly through the lens of autoregressive (AR) feature extraction and various machine learning classifiers. The primary objective was to develop a robust model capable of accurately distinguishing between ten distinct hand gestures, a task that is inherently complex due to the overlapping electrical activities of flexor and extensor muscles. The findings reveal both the strengths and limitations of the methodologies employed, offering a pathway for future research and practical applications. One of the most notable outcomes of this study was the superior performance of the Linear Discriminant Analysis (LDA) classifier, which achieved an accuracy of 80% at an autoregressive order of 12. This result underscores the effectiveness of LDA in handling the high-dimensional feature space generated by the AR modeling process. LDA's ability to maximize the separation between different classes while minimizing the variance within each class is particularly advantageous in the context of EMG signals, where subtle differences in muscle activation patterns can be critical for accurate classification. The confusion matrix further illustrated this point, showing that certain gestures, such as "CLO," were classified with perfect accuracy, indicating that the model can effectively capture and differentiate between distinct muscle activation patterns associated with specific hand movements. However, the results also highlighted significant challenges in classifying certain gestures, particularly "IND" and "THI," which exhibited misclassification rates of 30% and 46%, respectively. These findings suggest that while some gestures can be reliably classified, others pose challenges due to overlapping features and potentially similar muscle activation patterns. The lower performance in these classes may be attributed to several factors, including the complexity of the gestures themselves, the inherent variability in muscle activation among different individuals, and the limitations of the feature extraction process. This variability is particularly relevant given that the dataset comprised participants with diverse anatomical and physiological characteristics, which can influence the EMG signals recorded.

The sensitivity of the K-Nearest Neighbors (KNN) classifier to feature complexity was also evident, as it demonstrated a significant drop in performance, achieving only 50% accuracy at the same autoregressive order. This decline highlights a critical limitation of KNN, which relies heavily on the local structure of the data. As the dimensionality of the feature space increases, KNN can struggle to maintain its effectiveness, particularly in distinguishing between classes that may not be well-separated in

the high-dimensional space. This finding suggests that while KNN can be a useful tool in certain contexts, its application in complex multi-class problems like hand motion classification may require careful consideration of feature selection and dimensionality reduction techniques. The integration of autoregressive modeling with machine learning classifiers presents a promising avenue for enhancing the classification of EMG signals. The AR modeling process effectively transforms raw EMG data into meaningful features that capture temporal dependencies, thereby reducing noise and enhancing the essential patterns needed for classification. However, the results indicate that further refinement of the feature extraction process is necessary. Future research could explore different AR orders or incorporate additional signal features beyond autoregressive coefficients, such as time-domain and frequency-domain features, to improve classification accuracy. Moreover, the study's findings have broader implications for practical applications in fields such as prosthetics control, rehabilitation engineering, and human-computer interaction. The ability to accurately classify a wide range of hand motions can significantly enhance the functionality and responsiveness of prosthetic devices, allowing for more natural and intuitive control. Similarly, improved hand gesture recognition can facilitate more effective rehabilitation tools, enabling tailored therapy programs that adapt to individual patient needs.

While this study provides valuable insights into the classification of hand motions using Electromyography (EMG) signals, several limitations and weaknesses must be acknowledged. Firstly, the dataset utilized in this research, although comprehensive, consisted of only eight participants, which may not adequately represent the broader population. The limited sample size can introduce variability in the EMG signals due to individual differences in muscle anatomy, physiology, and activation patterns. This variability can affect the generalizability of the model, as the classifier may perform well on the training data but struggle with unseen data from different individuals. Additionally, the study focused on ten specific hand gestures, which, while providing a rich dataset, may not encompass the full range of hand motions encountered in practical applications. The complexity of classifying gestures that share similar muscle activation patterns, such as "IND" and "THI," highlights the challenge of distinguishing between closely related movements. This limitation suggests that the model may require further refinement and additional training data to improve its robustness in real-world scenarios. Another notable weakness is the reliance on autoregressive modeling for feature extraction. While AR modeling effectively captures temporal dependencies in the EMG signals, it may overlook other relevant features present in the data. Future research could benefit from integrating additional feature extraction techniques, such as time-frequency analysis or machine learning-based feature selection, to enhance the model's performance.

This study contributes to the growing body of research focused on the classification of hand motions using Electromyography (EMG) signals, particularly through the application of autoregressive (AR) feature extraction combined with various machine learning classifiers. When compared to similar studies in the field, several key differences and similarities emerge that highlight the unique contributions and potential areas for improvement in this research. For instance, previous studies have often focused on simpler hand gestures or binary classification tasks, which can limit the applicability of their findings to more complex scenarios. A study by Geethanjali successfully classified

a limited set of hand gestures using surface EMG signals, achieving high accuracy but only addressing a binary classification problem [17]. In contrast, this research aimed to classify ten distinct hand motions, thereby addressing a significant gap in the literature regarding multi-class classification of complex gestures [17]. The complexity of classifying multiple gestures simultaneously presents unique challenges, particularly in distinguishing between gestures that may share similar muscle activation patterns.

Moreover, while many studies have employed traditional machine learning classifiers such as Support Vector Machines (SVM) and K-Nearest Neighbors (KNN), this research's emphasis on Linear Discriminant Analysis (LDA) as the most effective classifier at an autoregressive order of 12 is noteworthy. Previous research by Raurale et al. demonstrated the effectiveness of LDA in classifying EMG signals, but their focus was primarily on fewer gestures and did not explore the potential of autoregressive modeling for feature extraction [25]. The integration of AR modeling in this study not only enhances the feature extraction process but also contributes to the robustness of the classification model. Additionally, the findings of this study align with those of other researchers who have highlighted the challenges of classifying complex gestures. For example, a study by Dwivedi reported similar difficulties in accurately classifying gestures with overlapping features, emphasizing the need for improved feature extraction techniques and more sophisticated classifiers [26]. This research echoes those sentiments, suggesting that while certain gestures can be reliably classified, others require further refinement in the feature extraction process and potentially the exploration of more advanced classification algorithms.

5 Conclusion

The aim of this study was to classify ten distinct hand motions using Electromyography (EMG) signals through autoregressive feature extraction, employing classifiers such as K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Linear Discriminant Analysis (LDA). Our findings revealed that LDA outperformed the other classifiers, achieving an accuracy of 0.8 at an autoregressive order of 12, while KNN exhibited a significant drop to approximately 0.5 at the same order, indicating its sensitivity to increased complexity in the feature set. The confusion matrix highlighted the classification performance, where "CLO" achieved perfect accuracy of 1.0, and other classes such as "RIN" and "THL" demonstrated strong performance, yet some classes like "IND" and "THI" revealed misclassification rates of 0.7 and lower, respectively. These insights suggest that while certain gestures can be reliably classified, others pose challenges due to overlapping features. For future work, it is essential to refine the feature extraction process, perhaps by exploring different AR orders or integrating additional signal features beyond autoregressive coefficients. Additionally, testing more complex classification algorithms could enhance performance for ambiguous gestures, particularly those with similar feature representations. Ultimately, improving accuracy and robustness in hand gesture recognition can significantly impact applications in prosthetics control, human-computer interaction, and rehabilitation engineering.

References

1. Fortune BC et al (2019) Low-cost active electromyography HardwareX, 6:e00085. https://doi.org/10.1016/j.ohx.2019.e00085
2. Kakoty NM, Hazarika SM (2011) Recognition of grasp types through principal components of DWT based EMG features. In: IEEE international conference on rehabilitation robotics. https://doi.org/10.1109/ICORR.2011.5975398
3. 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
4. Leonardis D et al (2015) An EMG-controlled robotic hand exoskeleton for bilateral rehabilitation. IEEE Trans Haptics 8(2):140–151. https://doi.org/10.1109/TOH.2015.2417570
5. Carrizosa E et al (2022) The tree based linear regression model for hierarchical categorical variables. Expert Syst Appl 203:117423. https://doi.org/10.1016/j.eswa.2022.117423
6. Hahne JM et al (2014) Linear and nonlinear regression techniques for simultaneous and proportional myoelectric control. IEEE Trans Neural Syst Rehabil Eng 22(2):269–279. https://doi.org/10.1109/TNSRE.2014.2305520
7. Liu L et al (2011) System identification of non-linear, dynamic EMG-Torque relationship about the elbow. In: IEEE 37th annual northeast bioengineering conference, System, pp 1–2
8. McDaid A et al (2014) Non-linear model-based control for the human-inspired robotic exoskeleton (HuREx) gait trainer. IFAC Proc 47(3):11830–11835. https://doi.org/10.3182/20140824-6-ZA-1003.00576
9. Bashar SK et al (2018) Developing a novel noise artifact detection algorithm for smartphone PPG signals: preliminary results. In: 2018 IEEE EMBS international conference on biomedical & health informatics (BHI). IEEE, pp 79–82. https://doi.org/10.1109/BHI.2018.8333374
10. Frigo C et al (2000) EMG signals detection and processing for on-line control of functional electrical stimulation. J Electromyogr Kinesiol 10(5):351–360. https://doi.org/10.1016/S1050-6411(00)00026-2
11. Hokajärvi I (2011) Electrode contact impedance and biopotential signal quality. Tampere University of Technology
12. Roy SS et al (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 2020, pp 1–5. https://doi.org/10.1109/NCETSTEA48365.2020.9119911
13. Ul Hassan Z et al. (2022) Electromyography and speech controlled prototype robotic car using CNN based classifier for EMG. In: 2022 international conference on emerging trends in electrical, control, and telecommunication engineering ETECTE 2022—Proceedings. https://doi.org/10.1109/ETECTE55893.2022.10007092
14. Triwiyanto T et al (2019) Investigation of the number of features and muscles for an effective hand motion classifier using electromyography signal. J Phys Conf Ser. https://doi.org/10.1088/1742-6596/1373/1/012051
15. Foroutannia A et al (2022) A deep learning strategy for EMG-based joint position prediction in hip exoskeleton assistive robots. Biomed Signal Process Control. 75:103557. https://doi.org/10.1016/j.bspc.2022.103557.
16. Wang L et al (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

17. Geethanjali P, Ray KK (2015) A low-cost real-time research platform for EMG pattern recognition-based prosthetic hand. IEEE/ASME Trans Mechatron 20(4):1948–1955. https://doi.org/10.1109/TMECH.2014.2360119
18. Tkach D et al (2010) Study of stability of time-domain features for electromyographic pattern recognition. J Neuroeng Rehabil 7:21. https://doi.org/10.1186/1743-0003-7-21
19. Tkach D et al (2010) Study of stability of time-domain features for electromyographic pattern recognition. J Neuroeng Rehabil 7, 1:21. https://doi.org/10.1186/1743-0003-7-21
20. Bi L et al (2019) A review on EMG-based motor intention prediction of continuous human upper limb motion for human-robot collaboration. Biomed Signal Process Control 51:113–127. https://doi.org/10.1016/j.bspc.2019.02.011
21. Triwiyanto T et al (2021) A review: sensory system, data processing, actuator type on a hand exoskeleton design. J. Biomimetics, biomater Biomed Eng 50:39–49. https://doi.org/10.4028/www.scientific.net/JBBBE.50.39
22. Halaki M, Ginn KA (2012) Normalization of EMG signals: to normalize or not to normalize and what to normalize to? Comput. Intell Electromyogr Anal—A Perspect Curr Appl Futur Challenges 175–194. https://doi.org/10.5772/49957
23. Hu XHX et al (2008) Classification of surface EMG signal based on energy spectra change. In: 2008 international conference on biomedical engineering and informatics. vol 2, pp 1299–1303. https://doi.org/10.1109/BMEI.2008.24
24. Ranniger CU, Akin DL (1997) EMG mean power frequency determination using wavelet analysis. In: Proceedings of the 19th annual international conference of the IEEE engineering in medicine and biology society.'magnificent milestones and emerging opportunities in medical engineering (Cat. No.97CH36136). 4, C, pp 1589–1592. https://doi.org/10.1109/IEMBS.1997.757017
25. 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
26. 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

Feature Evaluation for Elbow Joint Angle Prediction Using EMG: A Preliminary Investigation

Triwiyanto[1(✉)], Levana Forra Wakidi[1], I. Putu Alit Pawana[2], Sari Luthfiyah[1], Bedjo Utomo[1], and Vijay Anant Athavale[3]

[1] Department of Medical Electronics Technology, Poltekkes Kemenkes Surabaya, Surabaya, Indonesia
triwiyanto@ieee.com
[2] Rehabilitation Unit, Dr. Soetomo Hospital, Surabaya, Indonesia
[3] Walchand Institute of Technology, Solapur, India

Abstract. This study investigates the performance of three time-domain features—Mean Absolute Value (MAV), Variance (VAR), and Root Mean Square (RMS)—in predicting angular motion using a time series dataset. The main problem addressed is the fluctuation in prediction accuracy over time, which is often overlooked in other studies that primarily focus on aggregate metrics. The aim of the research is to analyze and compare the prediction errors of these features, while also exploring their linearity with a normalization variable ('nvar') to enhance feature selection for predictive modeling. The method involves examining prediction errors over a 50-s period using normalized angles, and comparing the features using boxplots, scatter plots, and time series error analysis. Results show that MAV had a median error near zero with some outliers, while VAR exhibited larger negative errors, with a median around −2. RMS, on the other hand, displayed minimal variability, with errors concentrated around zero. Over the time series, prediction errors fluctuated between − 5 and 5 for all features, revealing dynamic changes in model accuracy. The strong linear relationships observed between the features and 'nvar' suggest that normalization improves feature performance. In conclusion, while MAV, VAR, and RMS are effective predictors, their errors vary over time, highlighting the need for temporal error analysis in predictive modeling. These findings emphasize the importance of feature selection and normalization to enhance model accuracy, which ranged from 0 to 1 in this study.

Keywords: elbow joint · prediction · time domain feature · RMS

1 Introduction

The accurate prediction of joint angles is vital for developing effective rehabilitation therapies, especially for individuals recovering from strokes [6, 12, 23]. Stroke significantly impacts motor functions, leading to decreased mobility and difficulties in performing daily tasks [14]. Rehabilitation is critical in this context, as it aims to restore motor skills,

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 507–520, 2026.
https://doi.org/10.1007/978-981-95-9233-3_37

improve independence, and enhance quality of life. Traditional methods for measuring joint angles, such as goniometers or optical motion capture systems, are often cumbersome, invasive, and expensive [22]. These tools can restrict patient movement, creating discomfort during therapy sessions. Furthermore, their reliance on specialized equipment can hinder real-time feedback, which is essential for both patients and therapists in monitoring progress. In many rehabilitation scenarios, the need for immediate and accurate assessments of joint angles is paramount. Patients undergoing physical therapy require constant motivation and adaptive strategies to facilitate recovery, making the need for user-friendly and non-invasive monitoring systems even more pronounced. Current technologies may fail to meet these requirements, thereby limiting their application in everyday therapeutic settings. Additionally, the variability of patient responses to rehabilitation exercises necessitates a more individualized approach, where understanding joint angles in real-time could play a transformative role [17].

The integration of real-time data analytics into rehabilitation programs could enhance the personalization of therapy [15]. However, this integration is hindered by the limitations of existing measurement techniques. For instance, traditional goniometers require physical attachment to the limb, which may not be suitable for all patients, particularly those with significant motor impairment. Similarly, optical systems necessitate an unobstructed line of sight and complex setups, which are impractical in many clinical environments. Thus, there is a pressing need for more accessible, non-invasive techniques that not only predict joint angles but also provide instantaneous feedback. These systems should be adaptable, cost-effective, and user-friendly to facilitate better patient outcomes.

Moreover, the rehabilitation field increasingly recognizes the importance of technology-enhanced solutions, especially in light of advancements in wearable devices and sensor technologies. There is potential for integrating such technologies into therapeutic protocols, yet many existing studies remain focused on traditional methodologies. By exploring new avenues, such as using electromyography (EMG) for joint angle estimation, we can bridge the gap between technological advancement and practical application in rehabilitation settings [2, 24, 29]. However, achieving this requires rigorous investigation into the effectiveness of various features extracted from EMG signals, which remains an underexplored area in current research. Addressing these challenges and identifying optimal methods for joint angle prediction will significantly enhance rehabilitation strategies and ultimately improve patient care.

Recent advancements in electromyography (EMG) technology have significantly enhanced our ability to estimate joint angles non-invasively. EMG measures the electrical activity of muscles, providing valuable insights into muscle activation patterns during movement. Current state-of-the-art methods typically utilize various signal processing techniques alongside machine learning algorithms to analyze EMG signals. These methods allow for the derivation of joint positions based on the intricate relationship between muscle activity and joint movement. Researchers have explored a variety of features derived from EMG signals, including time-domain features (such as MAV, RMS, and VAR) [3, 25, 27], frequency-domain features [11, 19], and wavelet transforms [5, 18], among others .

In particular, machine learning techniques have gained traction due to their ability to model complex relationships within data. Algorithms such as support vector machines (SVM) [9], artificial neural networks (ANN) [16], and decision trees have been employed to predict joint angles based on EMG inputs [10]. These models often require substantial datasets for training, which can complicate their implementation in clinical settings where variability among patients is high. The complexity of these models can also introduce challenges in interpretability, making it difficult for clinicians to understand how predictions are derived.

However, many existing studies have focused on the application of these advanced machine learning techniques without thoroughly investigating the foundational aspect of feature selection. The choice of features used in the analysis can significantly impact the model's performance, yet the comparative effectiveness of various features, particularly for elbow joint angle prediction, remains underexplored. Features such as MAV and RMS have shown promise in previous studies, but their relative contributions to predicting elbow joint angles specifically need further examination.

Moreover, while machine learning provides powerful predictive capabilities, it is essential to establish a clear understanding of which features are most informative before delving into more complex modeling approaches. This foundational knowledge will aid in creating more efficient and interpretable models that can be easily applied in clinical practice. A robust understanding of feature extraction can streamline the process of model development and improve the accuracy of predictions, ultimately leading to better rehabilitation outcomes.

In addition to feature extraction, the integration of real-time data analytics into rehabilitation programs has emerged as a significant area of interest. The ability to monitor muscle activity and predict joint angles in real-time could facilitate immediate adjustments to therapy protocols, enhancing patient engagement and motivation. However, the existing literature often overlooks the practical aspects of implementing these systems in everyday clinical settings. Understanding the feasibility and effectiveness of different feature extraction methods in real-world applications is crucial for advancing this field.

To address these gaps, it is essential to conduct rigorous investigations into the most effective EMG features for predicting elbow joint angles. This preliminary study aims to focus on specific features—MAV, RMS, and VAR—to determine their predictive capabilities. By validating these predictions using the MPU6050 sensor, we aim to establish a reliable baseline for future research. Our exploration of feature extraction methods will not only contribute to the current understanding of joint angle prediction but also provide insights that could guide the development of more effective, user-friendly rehabilitation solutions.

2 Method

2.1 Dataset

The dataset for this study comprises EMG signals collected from participants during specific elbow joint movements, alongside corresponding joint angle measurements validated using the MPU6050 sensor. The aim was to create a comprehensive dataset that captures a range of muscle activation patterns associated with different elbow angles.

Participants were selected based on specific criteria, including age, physical condition, and prior experience with rehabilitation exercises. This selection ensured a diverse representation of muscle activation patterns, which is critical for evaluating the effectiveness of various EMG features.

To construct the dataset, we recorded EMG signals from the biceps and triceps muscles, as these are primarily responsible for elbow flexion and extension. The signals were captured using surface EMG electrodes placed on the skin over the targeted muscles, following standardized electrode placement guidelines to ensure consistency and accuracy in data collection. Each participant performed a series of movements, including flexion, extension, and isometric contractions at predetermined angles (e.g., 0°, 30°, 60°, 90°, and 120°). This range of angles was chosen to encompass the typical movement spectrum observed in daily activities and rehabilitation exercises.

In parallel, the elbow joint angles were measured in real-time using the MPU6050 sensor, a widely used device that combines a 3-axis gyroscope and a 3-axis accelerometer. This sensor was mounted on the participants' forearms, providing accurate angle measurements as they executed the movements. The EMG signals and joint angle data were synchronized using timestamps to ensure the integrity of the dataset. The final dataset consisted of multiple trials from each participant, allowing for a robust analysis of the relationships between EMG features and elbow joint angles.

2.2 Data Collection

Data collection was performed in a controlled environment, ensuring minimal external interference and consistent conditions across all trials. Before data collection commenced, each participant underwent a brief orientation session to familiarize them with the procedures and the importance of consistent effort during the movements. Participants were instructed to perform the elbow movements at a self-selected speed, allowing for natural muscle activation patterns while still maintaining control over their movements.

During each trial, EMG signals were recorded at a sampling rate of 1000 Hz to capture the rapid changes in muscle activity associated with dynamic movements [1, 4, 20, 28]. The data was collected using a multi-channel EMG system capable of processing signals from multiple electrodes simultaneously. In parallel, the MPU6050 sensor continuously recorded the elbow joint angle during the movements, with data synchronized to the EMG recordings via a common timing mechanism.

Following the completion of the trials, the raw data underwent preliminary checks to identify any artifacts or inconsistencies, such as noise or irregular signals. Any trials that exhibited significant interference (e.g., from equipment malfunction or movement artifacts) were excluded from the final dataset to ensure high data quality. This meticulous data collection process was crucial in ensuring that the dataset was both comprehensive and reliable for subsequent analyses.

2.3 Data Processing

Data processing involved several critical steps to prepare the raw EMG and joint angle data for analysis. The first step was filtering the raw EMG signals to remove noise and

artifacts. A bandpass filter with cutoff frequencies of 20 Hz to 450 Hz was applied, as this range captures the essential frequency components of muscle signals while filtering out low-frequency noise and high-frequency artifacts.

Once the signals were filtered, they were rectified to convert all values to positive, allowing for the computation of relevant features. Features of interest included Mean Absolute Value (MAV), Root Mean Square (RMS), and Variance (VAR), each of which provides different insights into the muscle activation patterns. MAV was calculated as the average of the absolute values of the EMG signal, while RMS provides a measure of the signal's power over a defined window. Variance offered insights into the signal's variability, which can reflect muscle recruitment strategies.

After calculating these features for each trial, the data was normalized to account for individual differences in muscle size and strength. Normalization was performed by dividing each feature by its maximum value across all trials for each participant, ensuring that the subsequent analysis focused on relative changes rather than absolute values.

The joint angle data obtained from the MPU6050 sensor was also processed to remove any noise and ensure accuracy. This involved applying complementary filtering techniques to fuse data from the gyroscope and accelerometer, resulting in a more stable measurement of elbow angles. The processed joint angle data was then synchronized with the EMG feature data, forming a comprehensive dataset for analysis.

2.4 Data Analysis

Data analysis aimed to explore the relationships between the processed EMG features and the corresponding elbow joint angles. A series of statistical analyses were conducted to assess the predictive capabilities of the selected features. Initially, exploratory data analysis was performed to visualize the distributions and correlations between EMG features and joint angles, utilizing scatter plots and correlation matrices.

Subsequently, multiple regression analysis was employed to evaluate the extent to which the EMG features could predict elbow joint angles. Regression models were constructed using a stepwise selection process to identify the most significant predictors among the MAV, RMS, and VAR features. This process involved assessing the contribution of each feature to the overall model fit, providing insights into which features were most informative for joint angle prediction.

Cross-validation techniques, such as k-fold cross-validation, were applied to ensure the robustness of the regression models. This involved partitioning the dataset into k subsets, training the model on $k - 1$ subsets, and testing it on the remaining subset. This approach mitigated the risk of overfitting and ensured that the models generalized well to unseen data.

Finally, the predictive performance of the models was evaluated using metrics such as R-squared values and root mean square error (RMSE), which provided insights into the accuracy of the predictions. These analyses ultimately aimed to establish a clear understanding of the relationship between EMG features and elbow joint angles, contributing to the identification of optimal feature extraction methods for future rehabilitation applications.

3 Result

The graph titled ** “Angle Prediction”** presents a detailed analysis of four different variables over a time span of 50 s. The x-axis represents **Time (s)**, ranging from 0 to 50 s, while the y-axis on the left side is labeled **“Normalized angle”** and ranges from 0 to 1. The right y-axis, labeled **“Normalized angle prediction”**, also shares the same range from 0 to 1. This normalization suggests that the angles have been scaled to a common standard, facilitating easier comparison. The graph features four distinct lines, each representing a different variable: **’nang’ (Red Line)**, **’nvar’ (Blue Line)**, **’nms’ (Green Line)**, and an unnamed purple line. The red line exhibits sharp peaks that nearly reach the maximum normalized value of 1, followed by steep declines towards the minimum value. The periodic nature of these peaks indicates a highly variable angle over time. The blue line shows less pronounced peaks compared to ‘nang’, suggesting lower angle values or less variation over time. The peaks and troughs are more moderate, indicating a more stable variable. The green line has more frequent oscillations than both ‘nang’ and ‘nvar’, with its peaks not reaching as high as ‘nang’. The frequent oscillations suggest a variable that changes more rapidly over time. Although not labeled in the legend, the unnamed purple line follows a pattern similar to ‘nms’, with slightly higher peak values. Its behavior suggests it might be another model or prediction closely related to ‘nms’. The graph likely represents data from an experiment or simulation involving angular motion or predictions. The periodic behavior of the lines suggests that the system being analyzed has regular cycles or oscillations. This could be relevant in various fields such as mechanical systems, signal processing, and control systems. The data might be from systems like pendulums, engines, or other mechanical devices where angles change over time. The graph could be used to assess the performance of prediction algorithms in signal processing applications. Engineers might use this data to design and evaluate control systems that need to predict and manage angular positions. Normalization of the angles is crucial as it allows for the comparison of different variables on the same scale. This is particularly useful when dealing with data that might have different units or magnitudes. By normalizing the data, the graph provides a clear visual representation of how each variable behaves relative to the others. Understanding the behavior of these variables over time can have several practical applications. In mechanical systems, predicting angular positions can help in maintenance and avoiding failures. In robotics, accurate angle prediction is essential for movement and task execution. In aerospace engineering, predicting angles can be critical for navigation and control of aircraft and spacecraft.

The **“Angle Prediction”** graph provides valuable insights into the behavior of four different variables over time. By normalizing the angles, it allows for easy comparison and analysis, making it a useful tool in various engineering and scientific applications. The periodic nature of the data suggests regular cycles, which could be leveraged in predictive maintenance, robotics, aerospace, and other fields where understanding and predicting angular motion is crucial.

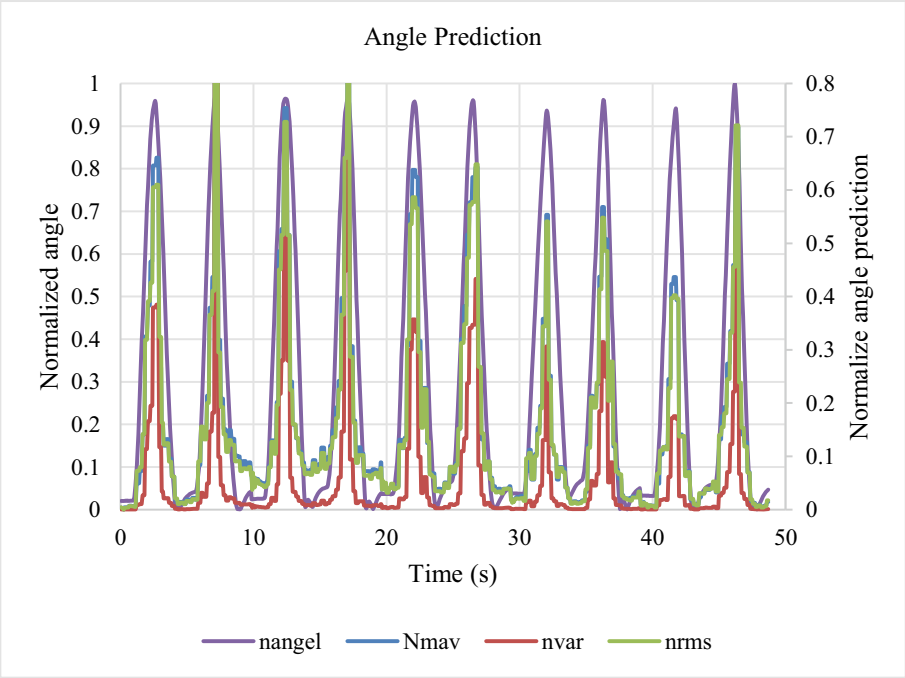


Fig. 1. Normalized angle predictions over time: comparative analysis of four variables.

Figure 1 titled "Error Distribution" presents a boxplot comparison of error distributions for three different features: MAV, VAR, and RMS. Each feature is represented by a distinct color—MAV in blue, VAR in red, and RMS in green. The vertical axis is labeled "Error" and ranges from − 5 to 5, indicating the magnitude of errors for each feature. For the MAV feature, the blue boxplot shows a median error close to 0, with the interquartile range (IQR) extending from approximately − 1 to 1. This suggests that the majority of MAV errors are relatively small, clustered around the median. However, there are several outliers indicated by individual dots above and below the IQR, showing that some MAV errors deviate significantly from the median. The VAR feature, represented by the red boxplot, has a median error around − 2, with an IQR stretching from about − 3 to − 1. This indicates that VAR errors tend to be more negative, with a central tendency lower than that of MAV. Similar to MAV, the VAR feature also has numerous outliers scattered above and below the main body of the boxplot, highlighting the presence of significant deviations in some instances. Lastly, the RMS feature is depicted by the green boxplot. Interestingly, both the lower and upper quartiles for RMS are very close to or equal to 0, resulting in a boxplot with no visible IQR or whiskers. This suggests that the RMS errors are highly concentrated around the median, with minimal variability. However, there is one outlier slightly above 0, indicating a rare instance of deviation from the central tendency. This boxplot graph provides a visual comparison of the error distributions for the three features, highlighting their central tendencies, variability, and the presence of outliers. Such a comparison is valuable in understanding the performance and reliability

of different features in a dataset or analysis context. Boxplots are particularly useful for identifying medians, ranges, and variability within data sets, as well as spotting outliers that could be potential anomalies or errors in measurement.

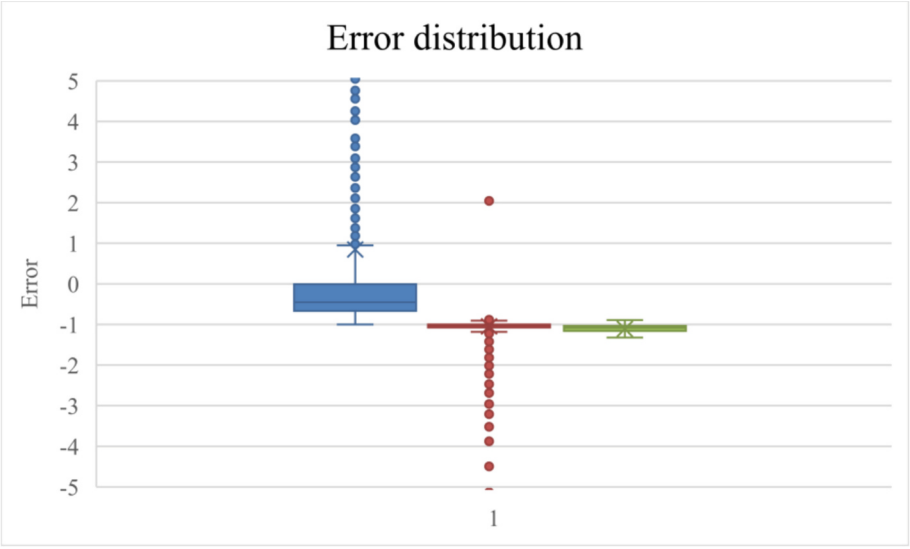


Fig. 2. Boxplot comparison of error distributions for MAV, VAR, and RMS features.

Figure 2 titled "Normalized Feature Linearity" presents a scatter plot that examines the linearity of three different features: MAV (Mean Absolute Value), VAR (Variance), and RMS (Root Mean Square). The x-axis is labeled 'nvar', and the y-axis is labeled 'Normalized feature'. Each feature is represented by a distinct color: green for MAV, red for VAR, and blue for RMS. The MAV data points, shown in green, are spread across the entire range of the x-axis but are mostly concentrated towards the higher end of the y-axis. This distribution suggests that MAV values tend to be higher when normalized, indicating a strong linear relationship with 'nvar'. The VAR data points, depicted in red, form a clear upward trend as 'nvar' increases. This indicates a proportional relationship between 'nvar' and VAR values, with higher 'nvar' values corresponding to higher normalized VAR values. The RMS data points, shown in blue, also exhibit an upward trend but start from the lower end of both axes. This suggests that RMS values increase proportionally with 'nvar', but the relationship starts from a lower baseline compared to MAV and VAR. This scatter plot visually represents how the different features—MAV, VAR, and RMS—scale with the normalization variable 'nvar'. The clear trends observed in the data points indicate that all three features have a linear relationship with 'nvar', but the strength and starting points of these relationships vary. MAV shows the highest concentration towards the upper end of the y-axis, indicating a stronger linearity, while RMS starts from a lower baseline, suggesting a more gradual increase. Understanding the linearity of these features is crucial in fields like signal processing or machine learning, where feature selection and engineering play a significant role in model performance.

The linear relationships depicted in this figure can help in selecting the most appropriate features for predictive modeling, ensuring that the models can accurately capture the underlying patterns in the data can be seen on Fig. 3.

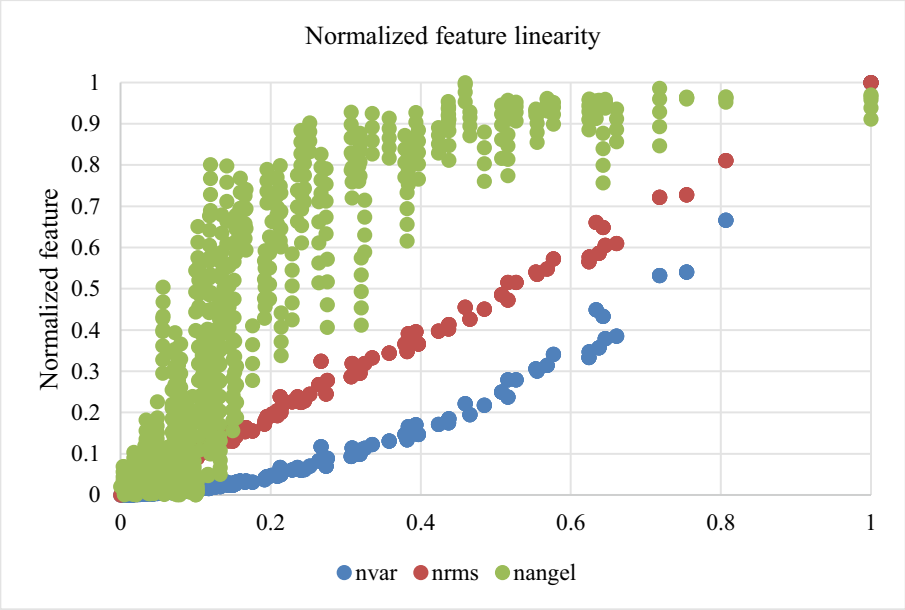


Fig. 3. Linearity analysis of normalized features: MAV, VAR, and RMS.

Figure 4 titled "Error in Prediction" presents a line graph that illustrates the prediction error over time for three different time domain features: Mean Absolute Value (MAV), Variance (VAR), and Root Mean Square (RMS). The x-axis represents time in seconds, ranging from 0 to 50, while the y-axis represents the prediction error, with values ranging from − 5 to 5. Three lines are plotted on the graph, each corresponding to one of the features. The blue line represents the error for MAV, the green line represents the error for VAR, and the red line represents the error for RMS. Each line fluctuates over time, showing the variations in prediction error for each feature. The blue line (Error mav) shows a series of peaks and troughs, indicating moments where the prediction error for MAV was higher or lower. The fluctuations suggest that the prediction model's accuracy varies significantly over time when using MAV as a feature. The green line (Errvar) also exhibits fluctuations, but with a different pattern compared to the MAV error. The peaks and troughs in the VAR error line indicate that the prediction error for VAR also varies over time, but the pattern may suggest different underlying dynamics or model behavior. The red line (Errrms) shows a similar fluctuating pattern, with peaks and troughs indicating the prediction error for RMS. The variations in the RMS error line suggest that the prediction model's accuracy also changes over time when using RMS as a feature. Overall, this graph provides a visual representation of how the prediction errors for MAV, VAR, and RMS change over time. The fluctuations in the error lines

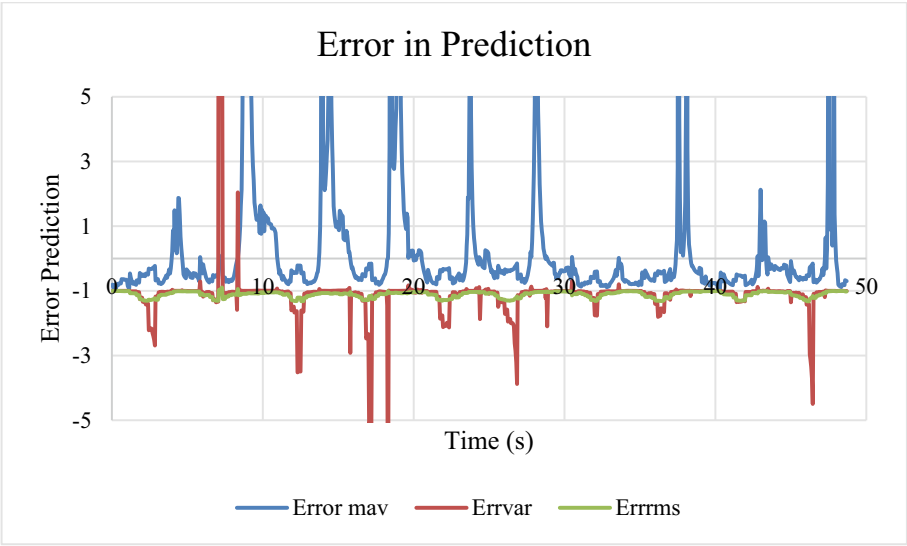


Fig. 4. Time series analysis of prediction errors for MAV, VAR, and RMS features.

highlight the dynamic nature of the prediction model's performance and suggest that different features may have varying levels of predictive accuracy at different times. This information can be valuable for analyzing the performance of the predictive model and identifying areas for improvement in feature selection or model tuning.

4 Discussion

In this study, we investigated the predictive performance of three time-domain features—Mean Absolute Value (MAV), Variance (VAR), and Root Mean Square (RMS)—in forecasting elbow joint angles based on EMG signals. The analysis of prediction errors revealed significant fluctuations over the 50-s observation period, indicating that the model's accuracy is not only variable but also sensitive to the choice of features used. The MAV feature, represented by the blue line in the error graph, exhibited pronounced peaks and troughs, suggesting that it is particularly responsive to transient changes in the EMG data. This sensitivity may be beneficial in capturing rapid muscle activation patterns; however, it also introduces a level of unpredictability in the model's performance. Such variability could be attributed to the inherent noise present in EMG signals, which MAV may amplify during periods of high muscle activity. This finding underscores the necessity for robust preprocessing techniques to mitigate noise and enhance the reliability of MAV as a predictive feature. In contrast, the VAR feature, depicted by the green line, displayed a different pattern of fluctuations. While it also experienced variability, the nature of these fluctuations suggests that VAR may provide a more stable measure of muscle activity over time. This stability could make VAR a more reliable feature for long-term predictions, particularly in clinical settings where consistent performance is crucial. However, the limitations of focusing solely on these three features must be

acknowledged. The analysis was constrained by a relatively short observation period of 50 s, which may not adequately capture longer-term trends and variations in muscle activity. Future research should consider extending the observation period to gain a more comprehensive understanding of the model's behavior. Moreover, the exploration of additional features beyond MAV, VAR, and RMS could significantly enhance predictive accuracy. Incorporating frequency domain features, wavelet transform features, and other statistical measures may provide deeper insights into the complex dynamics of muscle activation and improve the model's overall performance. By broadening the feature set, future studies can better account for the variability inherent in EMG signals and develop more robust predictive models that are applicable in diverse clinical scenarios. Ultimately, this research highlights the importance of feature selection and the need for ongoing refinement of predictive models in the context of EMG-based joint angle estimation.

In this study, we examined the predictive performance of three time-domain features—Mean Absolute Value (MAV), Variance (VAR), and Root Mean Square (RMS)—in estimating elbow joint angles from EMG signals. Our findings align with previous research that emphasizes the utility of these features in muscle activity analysis. For instance, a study by Triwiyanto et al. demonstrated that MAV is a robust indicator of muscle activation, particularly in dynamic movements, which supports our observation that MAV effectively captures transient changes in muscle activity [26]. However, the variability in MAV's prediction accuracy over time highlights its sensitivity to noise, a limitation also noted by Islam et al., who found that MAV can be influenced by external factors such as electrode placement and movement artifacts [8].

In contrast, VAR has been less frequently explored in the literature, despite its potential to provide insights into muscle activation variability. Our results indicate that VAR offers a more stable measure compared to MAV, which is consistent with findings from a study by Coyle et al., where VAR was utilized to assess muscle recruitment strategies during different tasks [13]. This stability may make VAR a more suitable feature for long-term predictions, particularly in clinical settings where consistent performance is essential. The normalization techniques employed in this study also merit discussion. Previous studies, such as those by De Luca, have highlighted the importance of normalizing EMG signals to account for individual differences in muscle size and strength, which can significantly impact predictive accuracy [7]. Our findings reinforce this notion, demonstrating that normalization can enhance the performance of selected features, thereby improving model reliability. While this study focused on a limited set of features, it opens avenues for future research to explore a broader range of features, including frequency domain and wavelet transform features. For example, a study by Wang et al. suggested that incorporating frequency domain features can provide additional insights into muscle dynamics and improve the accuracy of joint angle predictions [21]. Overall, this study contributes to the growing body of literature on EMG-based predictive modeling, emphasizing the importance of feature selection, normalization, and the exploration of additional features to enhance model robustness and applicability in clinical practice.

In this study, several limitations and weaknesses were identified that may impact the overall findings and their applicability to real-world scenarios. One significant limitation

is the relatively short duration of the data collection period, which spanned only 50 s. While this timeframe allows for a focused analysis of the model's performance, it may not adequately capture longer-term trends and variations in prediction accuracy. Extended observation periods could provide a more comprehensive understanding of how the predictive models behave over time, potentially revealing additional factors that influence prediction errors. Another notable weakness is the exclusive focus on three time-domain features: Mean Absolute Value (MAV), Variance (VAR), and Root Mean Square (RMS). Although these features are widely recognized for their effectiveness in analyzing electromyography (EMG) signals, they represent only a fraction of the potential features available for exploration. Future research should consider incorporating a broader range of features, including frequency domain features, wavelet transform features, and other statistical measures. This expansion could enhance the model's predictive accuracy and provide deeper insights into muscle activation patterns. Additionally, the study's reliance on a specific dataset may limit the generalizability of the findings. Variability in participant characteristics, such as age, gender, and physical condition, could influence the results. Therefore, future studies should aim to include a more diverse participant pool to validate the findings and ensure that the conclusions drawn are applicable across different populations. Addressing these limitations will be crucial for advancing the understanding of EMG features in predicting joint angles and improving rehabilitation strategies.

5 Conclusion

The aim of this study was to analyze the performance of three time-domain features—Mean Absolute Value (MAV), Variance (VAR), and Root Mean Square (RMS)—in predicting angular motion and assess their prediction errors over time. The findings show that each feature exhibited different levels of prediction accuracy and error variability. The MAV feature demonstrated a median error close to 0, with several outliers, while the VAR feature had a median error around -2, indicating larger negative errors. The RMS feature displayed minimal error variability, with most errors centered near zero. In the time series analysis, prediction errors fluctuated between -5 and 5 across all three features, reflecting dynamic changes in model accuracy over the 50-s period. The study also found strong linear relationships between the features and the normalization variable ('nvar'), suggesting that feature normalization plays a key role in improving predictive performance. Despite these results, the study has limitations, such as the short duration of 50 s, which may not capture long-term trends, and the focus on only three features. Future work could explore additional features and extend the analysis over a longer time period to provide a more comprehensive understanding of the model's behavior and further enhance prediction accuracy. Advanced normalization techniques could also be investigated to optimize feature selection and improve overall model performance in predictive tasks such as angular motion estimation.

References

1. Afsharipour B et al (2019) Two-dimensional surface EMG: the effects of electrode size, interelectrode distance and image truncation. Biomed Signal Process Control 49:298–307. https://doi.org/10.1016/J.BSPC.2018.12.001
2. Akhtar A et al (2017) Estimation of distal arm joint angles from EMG and shoulder orientation for transhumeral prostheses. J Electromyogr Kinesiol 35:86–94. https://doi.org/10.1016/j.jelekin.2017.06.001
3. Bhattacharya A et al (2017) Time domain multi-feature extraction and classification of human hand movements using surface EMG. In: The 2017 4th International conference on advanced computing and communication systems ICACCS 2017, pp 1–5. https://doi.org/10.1109/ICACCS.2017.8014594
4. Bilodeau M et al (1997) Normality and stationarity of EMG signals of elbow flexor muscles during ramp and step isometric contractions. J Electromyogr Kinesiol 7(2):87–96. https://doi.org/10.1016/S1050-6411(96)00024-7
5. Elamvazuthi I et al (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
6. Feys H et al (2018) An EMG-controlled robotic hand exoskeleton for bilateral rehabilitation. IEEE Trans Haptics 8(2):924–929. https://doi.org/10.3389/fneur.2017.00107
7. Florimond V (2010) Basics of surface electromyography applied to physical rehabilitation and. Thought Technol Ltd 1(March):1–50
8. Islam MJ et al (2023) Optimizing electrode positions on forearm to increase SNR and myoelectric pattern recognition performance. Eng Appl Artif Intell 122:106160. https://doi.org/10.1016/j.engappai.2023.106160
9. Kiani B et al (2021) Predictive torque control of induction motor drive with reduction of torque and flux ripple. Bull Polish Acad Sci Tech Sci 69(4):1–13. https://doi.org/10.24425/bpasts.2021.137727
10. Kim T et al (2021) Effects of sampling rate and window length on motion recognition using sEMG armband module. Int J Precis Eng Manuf 22(8):1401–1411. https://doi.org/10.1007/S12541-021-00546-6/METRICS
11. Lalitharatne TD et al (2013) Compensation of the effects of muscle fatigue on EMG-based control using fuzzy rules based scheme. Conf Proc IEEE Eng Med Biol Soc 2013:6949–6952. https://doi.org/10.1109/EMBC.2013.6611156
12. Leonardis D et al (2015) An EMG-controlled robotic hand exoskeleton for bilateral rehabilitation. IEEE Trans Haptics 8(2):140–151. https://doi.org/10.1109/TOH.2015.2417570
13. Lobo-Prat J et al (2014) Non-invasive control interfaces for intention detection in active movement-assistive devices. J Neuroeng Rehabil 11, 1:168. https://doi.org/10.1186/1743-0003-11-168
14. Loureiro RCV et al (2011) Advances in upper limb stroke rehabilitation: a technology push. Med Biol Eng Comput 49(10):1103–1118. https://doi.org/10.1007/S11517-011-0797-0
15. Lu J et al (2017) An in-field automatic wheat disease diagnosis system. https://doi.org/10.1016/j.compag.2017.09.012
16. Ngeo J et al (2012) Continuous estimation of finger joint angles using muscle activation inputs from surface EMG signals. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society (EMBS), pp 2756–2759. https://doi.org/10.1109/EMBC.2012.6346535
17. Nguyen T et al (2022) Improved exoskeleton human motion capture system. In: Khang NV et al (eds) Mechanisms and machine science. Springer Science and Business Media B.V, pp 669–679. https://doi.org/10.1007/978-3-030-91892-7_64

18. Phinyomark A et al (2012) Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification feature extraction and reduction of wavelet transform coefficients. https://doi.org/10.5755/j01.eee.122.6.1816
19. Phinyomark A et al (2012) Feature reduction and selection for EMG signal classification. Expert Syst Appl 39(8):7420–7431. https://doi.org/10.1016/j.eswa.2012.01.102
20. Phinyomark A, Scheme E (2018) A feature extraction issue for myoelectric control based on wearable EMG sensors. In: 2018 IEEE sensors applications symposium SAS 2018—Proceeding, 1–6 Jan 2018. https://doi.org/10.1109/SAS.2018.8336753
21. Shi WT et al (2018) A bionic hand controlled by hand gesture recognition based on surface EMG signals: a preliminary study. Biocybern Biomed Eng 38(1):126–135. https://doi.org/10.1016/j.bbe.2017.11.001
22. 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
23. Singh RM et al (2013) A review on surface EMG based control schemes of exoskeleton robot in stroke rehabilitation. In: 2013 international conference on machine intelligence and research advancement. IEEE, Katra, India, pp 310–315. https://doi.org/10.1109/ICMIRA.2013.65
24. Tang Z et al (2022) Effect of velocity and acceleration in joint angle estimation for an EMG-Based upper-limb exoskeleton control. Comput Biol Med 141:105156. https://doi.org/10.1016/j.compbiomed.2021.105156
25. Triwiyanto et al (2017) An investigation into time domain features of surface electromyography to estimate the elbow joint angle. Adv Electr Electron Eng 15(3):448–458. https://doi.org/10.15598/aeee.v15i3.2177
26. Triwiyanto et al (2016) Quantitative relationship between feature extraction of sEMG and upper limb elbow joint angle. In: 2016 international seminar on application for Technology of Information and Communication. IEEE, Surabaya, pp 44–50. https://doi.org/10.1109/ISEMANTIC.2016.7873808
27. Triwiyanto T et al (2019) Investigation of the number of features and muscles for an effective hand motion classifier using electromyography signal. J Phys Conf Ser. https://doi.org/10.1088/1742-6596/1373/1/012051
28. Triwiyanto T et al (2023) Single Lead EMG signal to control an upper limb exoskeleton using embedded machine learning on raspberry pi. J Robot Control 4(1):35–45. https://doi.org/10.18196/jrc.v4i1.17364
29. Wahyunggoro O et al. Quantitative relationship between feature extraction of sEMG and upper limb elbow joint angle

ECG Digital Filter Processing: A Review

Triwiyanto[1(✉)], Triana Rahmawati[1], Endang Dian Setyoningsih[1], and Vugar Abdullayev[2]

[1] Department of Medical Electronics Technology, Poltekkes Kemenkes Surabaya, Surabaya, Indonesia
triwiyanto@ieee.org

[2] Department of Computer Engineering, Azerbaijan State Oil and Industry University, Baku, Azerbaijan

Abstract. Electrocardiogram (ECG) signals are vital diagnostic tools used widely in clinical and research settings to monitor cardiac health. However, the quality and interpretability of ECG signals are often compromised by various types of noise and artifacts, such as power line interference, baseline wander, and motion artifacts. Effective removal of these unwanted disturbances is crucial to ensure accurate diagnosis and reliable analysis. This review paper provides a comprehensive overview of digital filtering methods employed in ECG signal processing, emphasizing the development and application of Infinite Impulse Response (IIR) and Finite Impulse Response (FIR) filters. The study discusses the characteristics, advantages, and limitations of common digital filters such as Butterworth, Chebyshev, and Elliptic filters, highlighting their roles in attenuating noise while preserving essential morphological features of the ECG waveform. It underscores the trend toward utilizing IIR filters due to their efficiency in achieving steep frequency responses with fewer coefficients, thus reducing computational load and enabling real-time processing. Additionally, the review examines the integration of digital filters within matrix systems like MATLAB, illustrating their widespread adoption for ECG data analysis. Critical challenges encountered in ECG filtering are addressed, including the trade-off between filter complexity and signal integrity, as well as potential artifacts introduced by filtering techniques. Furthermore, the paper explores emerging digital filtering algorithms and their potential enhancements in noise reduction, along with future research directions involving internet-based alarm systems and advanced adaptive filtering solutions for improved heart attack prevention and cardiac diagnostics. This review concludes by emphasizing the importance of continuous innovation in filter design and implementation to enhance the accuracy, reliability, and efficiency of ECG signal processing in biomedical applications.

Keywords: ECG · digital filter · noise · FIR · IIR

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 521–539, 2026.
https://doi.org/10.1007/978-981-95-9233-3_38

1 Introduction

Heart disease remains one of the leading causes of mortality worldwide and represents a major global health challenge. Recent reports indicate that approximately 60% of patients with cardiovascular conditions succumb to the disease. Electrocardiogram (ECG) signals play a critical role in the monitoring and diagnosis of cardiac health, providing essential information for the detection and management of various heart abnormalities [33]. ECG is a key important signal for coronary heart disease that focuses on cardiac electrophysiology [22]. The fundamental electrocardiogram (ECG) signal occupies a relatively low-frequency spectrum, typically ranging from 0.5 Hz to 100 Hz. This frequency range encompasses the essential morphological components of the cardiac cycle, including the P-wave, QRS complex, and T-wave, which are critical for clinical interpretation. Frequencies below 0.5 Hz are generally associated with baseline wander, often caused by respiration or patient movement, while frequencies above 100 Hz are usually attributed to high-frequency noise or muscle artifacts rather than true cardiac activity. Defining this bandwidth is particularly important for the design of acquisition systems and digital filters in biomedical instrumentation, ensuring that clinically relevant information is preserved while extraneous noise is effectively suppressed. Consequently, appropriate selection of the passband is a fundamental step in ECG signal processing for accurate monitoring, analysis, and diagnosis [12, 20, 32]. Electrocardiogram (ECG) signals are highly susceptible to various sources of noise that can obscure clinically relevant information and compromise diagnostic accuracy. Common disturbances include power line interference, baseline wander caused by respiration or electrode movement, motion artifacts due to patient activity, and electromagnetic interference from surrounding electronic devices. Effective suppression of these noise components is therefore essential to ensure reliable ECG acquisition and accurate clinical interpretation [1]. Power line interference (PLI) in ECG recordings arises primarily from inductive and capacitive coupling with surrounding electrical power sources during signal acquisition. This interference manifests as narrowband noise centered at 50 or 60 Hz, typically with a bandwidth of less than 1 Hz and an amplitude that may reach up to 50% of the peak-to-peak ECG signal. Eliminating such interference is crucial, as it can significantly obscure clinically relevant features of the ECG. Biomedical signal processing, therefore, places considerable emphasis on filtering techniques, which are now routinely integrated into modern ECG systems to support accurate analysis and interpretation [37, 48, 50]. The objectives of ECG signal processing are multifaceted, encompassing efforts to enhance measurement accuracy and reproducibility compared to manual interpretation, as well as to extract clinically relevant information that cannot be readily obtained through visual inspection alone. In many cases, ECG recordings are acquired in outpatient or ambulatory settings, where the signals are frequently contaminated by various sources of noise, including interference from other physiological processes (Fig. 1). Consequently, effective noise reduction constitutes a critical goal of ECG signal processing to ensure reliable diagnostic outcomes [8, 21].

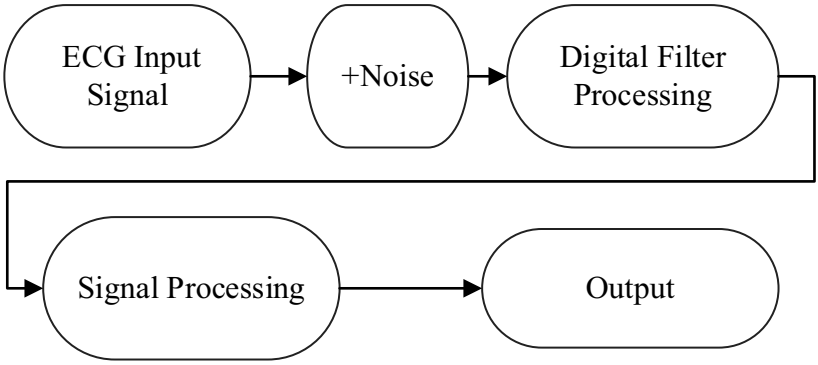


Fig. 1. ECG digital signal processing basic concept.

The ECG signal is disturbed by noise which has a frequency value of 50 Hz. Noise that interferes with the baseline of the ECG Signal will be removed. Then the data is given to the Stop Band Filter. That will eliminate interference of the power line to the ECG signal. Then the signal is passed through a Low Pass Filter with a frequency cut of 100 Hz. It will remove the remaining noise in the ECG signal. The original frequency of the ECG signal is between 0.5 Hz to 100 Hz. In principle, it is generally to eliminate noise from the electrocardiogram (ECG) using an IIR filter [43, 45]. Due to the noise interference of Baseline, it becomes difficult to research ECG records either manually or in an automated way. Different types of noise such as baseline roaming, power line interference will affect the original ECG signal. It will be reduced and eliminated by using this filtering technique. Many people today lose their lives because of liver attacks also known as myocardial infarctions [38]. This happens when blood circulation stops through the heart. Salah satu masalah terbesar dengan rekaman EKG ini adalah artefak dan suara. Various types of interference faced by ECG signals include electromyogram (EMG) interference, power line interference and movement artifacts that can cause poor signal quality. The movement of artifacts can produce deviations in the ECG recording of the signal [2, 23, 24]. The purpose of this study is to discuss the use of digital filters in ECG machines because it has been proven in several studies that the use of analog filters has a high cost and requires a large series to support an ECG machine. So at this time many studies discuss the use of digital filters in ECG machines because they are more cost-effective and do not require a large space for a series of filters and are more effective because there are several order options that can be used simultaneously.

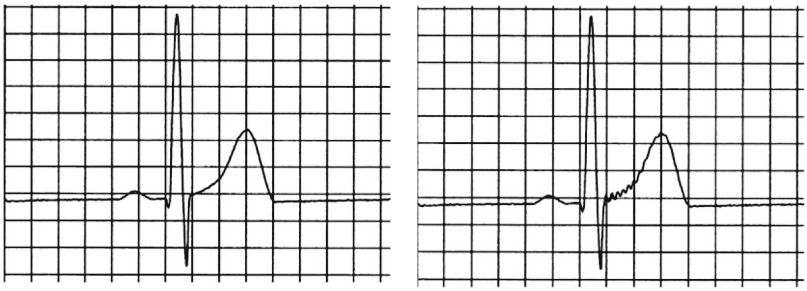


Fig. 2. Notch filters and ringing artifacts. Input (a) is the ANE20000 analytical wave per IEC 60601-2-51. The output of the notch filter (b) indicates the ringing after the QRS complex. This ringing is caused by a sudden change in the frequency domain spectrum.

2 Digital Filter Method

Electrical grid interference represents a prevalent source of noise contamination in medical electronic systems, particularly in diagnostic instrumentation such as electrocardiographs (ECGs). Early analog data acquisition systems incorporated common-mode rejection circuitry (Fig. 2) to mitigate power line frequency interference, typically at 50 or 60 Hz. Despite these design efforts, residual channel-specific interference voltages are still observed in clinical ECG recordings. To ensure high-fidelity signal acquisition, bandpass filters (BPFs) are routinely employed throughout the recording process. A BPF selectively transmits signals within a defined frequency range bounded by lower and upper cutoff frequencies (fc_1 and fc_2) while attenuating components outside this range, typically at the $-$ 3 dB points. When the passband is narrowly centered around the interference frequency, the filter functions as a notch filter, effectively suppressing power line artifacts. Adaptive notch filtering techniques may also be utilized to dynamically track and eliminate time-varying interference.

A critical issue associated with low-frequency filtering (LFF) in ECG devices is the introduction of ringing artifacts and oscillatory behavior following the QRS complexes [52]. From a signal processing perspective, sudden band-stop operations in the frequency domain often result in ringing effects, which manifest in the time domain as overshoot and undershoot near sharp transitions or edges. When the amplitude of these artifacts is pronounced, they can lead to clinically significant distortions, such as the appearance of artificial R waves, as well as ambiguous identification of J points and ST segments, thereby compromising the accuracy of detection and measurement [30, 40].

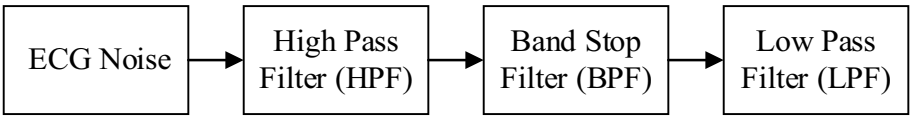


Fig. 3. Proses filter ECG.

In the ECG filtering process (Fig. 3), the acquired noisy ECG signal is first passed through a high-pass filter with a cutoff frequency of 0.5 Hz to suppress baseline wander and other low-frequency components. Subsequently, the signal is processed using a stopband (notch) filter with a narrow cutoff range between 59.5 Hz and 60.5 Hz, effectively eliminating power line interference. The filtered signal is then applied to a low-pass filter with a cutoff frequency of 100 Hz to remove high-frequency noise while preserving the clinically relevant components of the ECG. Since the useful frequency spectrum of the ECG signal lies between 0.5 Hz and 100 Hz, this filtering sequence ensures retention of diagnostic features such as the P-wave, QRS complex, and T-wave. Digital filters are preferred over analog filters because they provide superior performance with sharper transition bands, greater stability, and enhanced flexibility in adjusting filter parameters. Among various digital filter types, the Infinite Impulse Response (IIR) filter is widely used in digital signal processing (DSP) applications for ECG systems due to its computational efficiency and ability to achieve steep frequency responses with relatively few coefficients [29]. The advantage of IIR filters is that they require less coefficient for steep frequency responses, which can reduce the amount of computation time. The Butterworth configuration is one of the standard configurations of recursive filters in both analog and digital forms [5, 25, 31, 47]. This configuration emphasizes the approximation of low-pass filter characteristics, ensuring a smooth and gradual response approaching the zero point. A digital filter is fundamentally a computational algorithm that transforms an input sequence [] into an output sequence []. In this context, computation refers to the realization of mathematical operations such as integration, differentiation, and estimation. Digital filters can generally be designed using two approaches: the finite impulse response (FIR) and the infinite impulse response (IIR) methods. IIR filters provide efficient performance with fewer coefficients; however, the coefficients are typically non-integer, which may limit their suitability for certain real-time applications. In contrast, digital filters with simple integer coefficients are particularly advantageous for real-time systems, as they enable faster processing and simpler implementation. Such filters are characterized by efficient multiplication operations and the use of basic arithmetic units with only a limited number of shift and addition operations. For ECG signal conditioning, where real-time signal processing is critical, the filter design must achieve low computational complexity, rapid execution, linear phase behavior, and minimal signal distortion. The corresponding equations for the IIR filter are provided in Eqs. (1 and 2).

$$y(n) = \sum_{k=0}^{\infty} h(k)x(n-k) \tag{1}$$

$$y(n) = \sum_{k=0}^{N} bx(n-k) - \sum_{k=0}^{M} ax(n-k) \tag{2}$$

where h(k) is the impulse response of the filter which is theoretically unlimited in duration, while b and a are the filter coefficients(n) and y(n) are the inputs and outputs to the filter (Eq. (3)).

$$H(Z) = \frac{b0 \pm \mathrm{b1z} - 1 + \cdots + \mathrm{bNz} - \mathrm{N}}{a0 + a1z - 1 + \cdots + aMz_M} = \frac{\sum_{k=0}^{N} bkz - k}{\sum_{k=0}^{M} akz - k} \tag{3}$$

It is very important to find the appropriate values for the coefficients b and a. For the current output y(n) is a function of the previous output y(n-k) [53]. So this shows some kind of feedback system. The primary advantage of the infinite impulse response (IIR) filter lies in its feedback mechanism, which provides flexibility in design and allows the transfer function to be represented in terms of poles and zeros. While the IIR filter structure is computationally efficient, it presents significant challenges in achieving a strictly linear phase response. In this study, the IIR structure is employed to approximate a finite impulse response (FIR) filter with simple integer coefficients, thereby satisfying the requirements for real-time ECG signal processing. The design methodology involves first developing a band-pass filter, followed by the design of an all-pass filter with equivalent delay and gain characteristics. The all-pass filter is then combined with the band-pass filter to compensate for phase distortion, ultimately yielding the desired filter performance [14].

3 Method

In this review article, we focus on ECG signal processing methods using digital filters. Articles are obtained from a variety of available databases, including IEEE explore, and the Science Direct database. The articles collected are in recent years. After applying the selection criteria, we got several articles that discussed in depth about ECG signal processing using digital filters. In the discussion of the state of the art, the author then specifies the articles that have been collected with the Dowsinnes criteria. After doing these specifications, the number of articles that are suitable for review is 40 articles out of 45 downloaded articles.

3.1 Trendline Articles

At this stage of the review, articles were systematically searched from the publication period spanning 2010 to 2022. The search was conducted using the Publish or Perish (PoP) application, which employs the Google Scholar search engine as its primary database. This approach was selected due to its broad coverage of peer-reviewed literature and its ability to capture both journal and conference publications relevant to biomedical engineering and medical electronics. The retrieved data were subsequently analyzed to identify publication trends over the specified period. The resulting trendline, as illustrated in Fig. 4, provides a clear visualization of the growth and research activity in the field of ECG signal processing with digital filters, highlighting increasing scholarly interest and advancements in this domain.

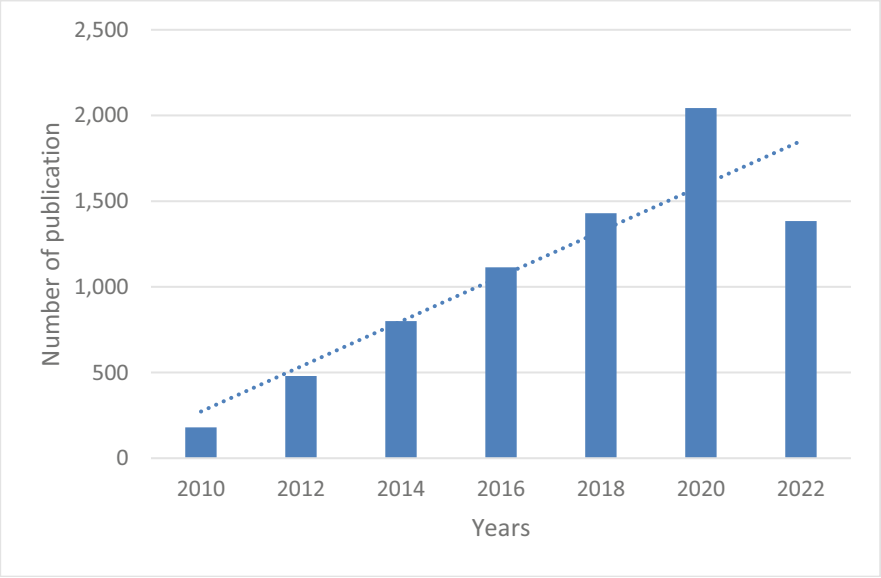


Fig. 4. Trendline publication by years.

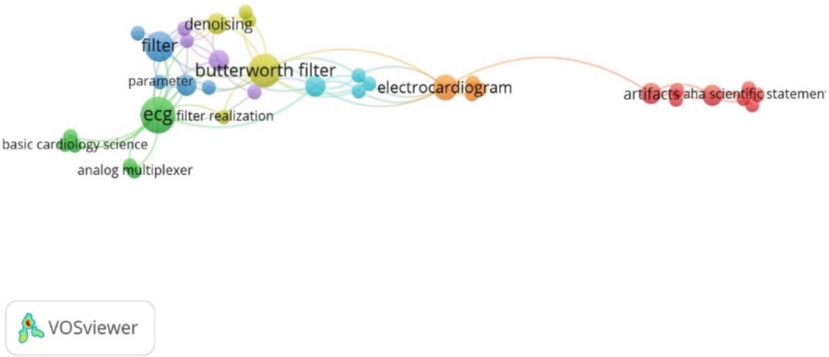


Fig. 5. Bibliographic analysis.

3.2 Keyword Analysis

Figure 5 presents the results of keyword mapping analysis conducted using the VOSviewer application. This visualization highlights the relationships and co-occurrence patterns among keywords extracted from the reviewed articles, thereby illustrating the primary research themes within the field of ECG signal processing using digital filters. The network clustering reveals distinct groups of frequently associated terms, indicating the dominant methodological approaches and areas of focus in the

literature, such as filter design, denoising strategies, and biomedical signal interpretation. To complement the visualization, a detailed description of the clustering results is provided in Table 1, which categorizes the keywords into thematic groups, offering a structured overview of the research landscape. Together, the figure and table facilitate a clearer understanding of how current studies are distributed across key topics and contribute to the development of state-of-the-art ECG signal processing techniques.

Table 1. Results of clustering keywords using VOSviewer

Cluster	Keywords	Color
1	Filter, parameter, electrocardiogram	Blue
2	Artifact, Aha scientific statemen	Red
3	ECG, analog multiplexer, basic cardiology science	Gree
4	Butterworth filter, denoising, ECG filter realization	Yellow

In the bibliometric analysis, all keywords in 50 articles were used with a minimum value of 2. The results show that there are 4 clusters that represent the relevance of keywords in one article to another. based on Fig. 1, ECG is strongly related to filters, Butterworth, and IIR.

4 Result

A comprehensive literature review has identified a range of digital filter designs implemented in electrocardiogram (ECG) devices. This review aims to support researchers and developers in advancing ECG instrumentation by synthesizing key methodologies and design approaches. Table 2 presents a summary of the filtering techniques and system architectures employed in recent ECG device development. The literature was sourced from peer-reviewed international journals and conference proceedings published over the past several years. In the subsequent sections, we highlight and evaluate the most effective filtering strategies, drawing on multiple authoritative sources.

Table 2. State of the art method of the Electrocardiograph

Authors	Method	Filter/Analytic method
Akanksha Mittal, 2015 [33]	The performance of various filtering techniques was evaluated by analyzing their impact on ECG waveform fidelity, with particular attention to power line interference at 50 Hz. Among the methods tested, the Chebyshev Window-based filter demonstrated superior results. Quantitative comparison of the Signal-to-Noise Ratio (SNR) between the filtered ECG signal and the original corrupted signal revealed that the Chebyshev Window effectively attenuated a substantial portion of the 50 Hz noise component. This confirms its efficacy in enhancing signal quality for clinical ECG applications	FIR/Hamming, Kaiser, and Chebyshev
Muhammad Abrar, 2018 [1]	This study presents an algorithm designed to suppress high-frequency noise in electrocardiogram (ECG) signals using a moving average filter as the primary denoising technique. Following the initial filtering stage, a polynomial curve fitting algorithm is applied to further smooth the ECG waveform, enhancing signal clarity and preserving morphological features critical for clinical interpretation.	Filter Notch & Wavelet/ SNR/ PSD

(*continued*)

Table 2. (*continued*)

Authors	Method	Filter/Analytic method
Nalini Singh, 2013 [46]	The ECG waveform utilized in this study was sourced from the MIT-BIH Arrhythmia Database, a widely recognized benchmark for evaluating signal processing algorithms in clinical cardiology. MATLAB was employed to implement the mathematical design of an Infinite Impulse Response (IIR) Butterworth filter, aimed at attenuating undesired frequency components. The clinical ECG data were sampled at a frequency of 360 Hz over a duration exceeding 10 s, ensuring sufficient temporal resolution for accurate signal analysis and filter performance evaluation.	IIR butterworth/FFT Spectrum
Xunyu Zhang, 2021 [54]	An experiment was conducted to demonstrate the effectiveness of discrete Fourier transforms (DFT) and Butterworth filters in eliminating high-frequency noise from electrocardiogram (ECG) signals. The study highlights the practical application of frequency-domain analysis and filter design in biomedical signal processing, emphasizing their role in enhancing ECG signal clarity for clinical interpretation.	Filter Butterworth/FFT

(*continued*)

Table 2. (*continued*)

Authors	Method	Filter/Analytic method
Y. A. Altay, 2019 [4]	The filter's effectiveness within the specified passband is evaluated both qualitatively and quantitatively using identical signal and noise model parameters under a consistent sampling frequency. Quantitative assessment is performed by comparing input and output signal-to-noise ratios (SNR) and calculating the root-mean-square error (MSE), providing objective metrics for denoising performance..	Filter Butterworth/SNR MSE
Kaustubh Manik, 2015 [10]	A fifth-order filter was designed to target the critical frequency range of ECG signals, spanning 0.5 Hz to 100 Hz. A low-pass cutoff frequency of 100 Hz was selected, with a passband attenuation of 1 dB to preserve signal integrity. Post-filtering, power spectral density (PSD) analysis was performed to evaluate signal quality..	Filter digital Butterworth, Chebyshev I, Chebyshev II, Elliptic orde 5 low pass dari frekuensi cutoff 100 Hz/ PSD
I Dewa Gede Hari Wisana, 2021 [39]	In this study, researchers evaluated filter performance by measuring attenuation levels in decibels (dB) across a frequency range of 0.5 Hz to 120 Hz. A function generator provided input to an Arduino analog channel embedded with a digital filter. Responses were observed at selected frequencies using 460 digital data points. The validated filter will be integrated into an ECG module.	IIR Butterworth, Chebyshev I, Chebyshev II, dan Elliptic; orde 2, 4, 6, 6, dan 10/

(*continued*)

Table 2. (*continued*)

Authors	Method	Filter/Analytic method
Ms. Chhavi 2018 [44]	ECG signals were retrieved from a database and imported into MATLAB for processing. These signals were then combined with simulated waveforms and noise components to emulate realistic conditions. The composite signals were analyzed in the time domain to evaluate filter performance and determine optimal design parameters for various digital filtering techniques.	FIR & IIR/Estimated weight convergence, Estimated curve, Comparison of Filter weights and estimated weights, Error curve
Faiz, Mohammed 2022 [13]	This paper presents a four-stage adaptive noise cancellation framework to suppress key ECG signal artifacts: baseline wander, motion artifacts, muscle noise, and 60 Hz power line interference. The study evaluates eight adaptive algorithms—including LMS, LMF, LMMN, SRLMS, SELMS, SSLMS, SRLMF, and SRLMMN based on their effectiveness in improving signal-to-noise ratio (SNR) during artifact removal.	LMS, LMF, LMMN, SRLMS, SELMS, SLMS, SRLMF,. SRLMMN

(*continued*)

Table 2. (*continued*)

Authors	Method	Filter/Analytic method
D. Preethi, 2018 [36]	Finite Impulse Response (FIR) filtering is a fundamental operation in digital signal processing (DSP), widely applied in biomedical systems. Conventional FIR filter designs using array multipliers often suffer from increased delay and power consumption. The proposed methodology introduces a digital architecture capable of reconfiguring FIR filters for high-performance, low-power medical electronics applications.	FIR notch filter
Thi Lan Nhi Pham, 2022 [26]	ECG signals were obtained from the MIT-BIH Arrhythmia Database for processing. The target application is a low-cost electrocardiogram device suitable for deployment in underserved hospitals. Baseline wander, typically below 0.5 Hz, was removed using MATLAB's Detrend function. Power line interference at 50 or 60 Hz was suppressed using a bandstop IIR filter.	IIR bandstop
Redmond, 2008 [41]	Four signal masks rail contact mask, high-frequency (HF) mask, low-power mask, and a final combined mask—are employed to detect key anomalies in ECG signals. These include amplifier saturation, HF artifacts, low signal amplitude, and the longest usable continuous ECG segment, ensuring reliable signal quality for clinical and diagnostic applications.	Notch filter by elliptic

(*continued*)

Table 2. (*continued*)

Authors	Method	Filter/Analytic method
Li, 2009 [27]	Key hemodynamic parameters analyzed include systolic blood pressure (SBP), diastolic blood pressure (DBP), and mean blood pressure (MBP), along with maximum positive and negative pressure slopes. Additional metrics comprise the duration of the maximum upward slope, time above the pulse-to-pulse interval (T) threshold, pulse pressure, and the delay time between ECG and arterial blood pressure (ABP) signals.	Modified KF
Mohammed Tali, 2019 [3]	Several Finite Impulse Response (FIR) filter models both low-pass and high-pass were analyzed to evaluate their performance in terms of response time, gain, harmonic distortion, and frequency rejection. The objective was to identify the optimal band-pass filter configuration capable of accurately reproducing ECG signals that closely reflect the patient's true cardiac activity.	FIR
Chien-Chen Tseng, 2019 [53]	The Chebyshev polynomial-based digital notch filter is an approximation technique used to achieve linear-phase design in IIR notch filters. Its key features include: (1) closed-form filter coefficients without requiring optimization, (2) precise zero-gain response at the specified notch frequency, and (3) adjustable 3-dB rejection bandwidth through filter order variation.	Digital notch filter

5 Discussion

Table 2 presents a state-of-the-art review of electrocardiography methodologies, compiled from recent journal publications and ongoing research in the field. The selected references primarily focus on the design of ECG systems incorporating digital filtering and signal analysis techniques. Various filter types such as low-pass, high-pass, bandpass, and notch filters are commonly employed in ECG circuits to enhance signal fidelity and suppress noise artifacts..

During the review process, the authors identified various applications of digital filters across biomedical systems. However, the predominant use remains in electrocardiogram (ECG) signal processing, where digital filtering techniques are essential for accurate analysis and clinical interpretation [51]. The objectives of ECG signal processing are multifaceted, encompassing enhanced measurement accuracy and reproducibility relative to manual interpretation, as well as the extraction of clinically relevant features not discernible through visual inspection alone. A critical function is noise reduction, as essential waveform components may be obscured and only detectable following the application of appropriate signal processing techniques [29, 47]. Electrocardiogram (ECG) signals are often contaminated by various types of noise, including power line interference (PLI), baseline wander, motion artifacts, and electromagnetic disturbances. To ensure accurate signal interpretation, ECG signals must undergo both analog and digital processing. In general, the standard sequence of ECG signal conditioning includes initial amplification using a low-noise instrumentation amplifier, followed by bandpass filtering to preserve the clinically relevant frequency range of the ECG. A notch filter, typically centered at 50 Hz (or 60 Hz depending on the region), is applied to suppress power line interference. Subsequently, a post-amplifier is employed to adjust the signal amplitude to levels compatible with the input range of the analog-to-digital converter (ADC) in the digital processing stage. It is worth noting that different studies may adopt varying cutoff frequencies for the applied bandpass filter, depending on the intended clinical or research application [7]. However, it was shown that the most commonly used low-frequency cutoff value was 0.5 Hz and the high frequency was between 40 Hz and 100 Hz compared to other bandpass cutoffs [35]. Compared to FIR filters, IIR filter implementations require fewer parameters, so less memory is used [6]. The use of IIR digital filters is widely used as noise reduction in ECG signals. The type of filter used is such as, butterworth filter [9], Chebyshev I, Chebyshev II and Elliptic [15–17].

A review of several studies indicates that infinite impulse response (IIR) filters offer the advantage of achieving steep frequency responses with fewer coefficients compared to finite impulse response (FIR) filters. This characteristic reduces the computational burden, thereby minimizing processing time and improving efficiency in real-time biomedical signal processing applications [28]. The Butterworth configuration is one of the standard configurations of recursive filters in both analog and digital forms [19]. Thus, digital filters can save the use of circuits on PCBs because digital filters do not require components such as analog filters, thus saving space and cost [11, 18, 34]. There is the use of the AD620 as a pre-amp where the AD620 has very unique characteristics and can be used as an instrumentation circuit on ECGs. In multi-lead ECG acquisition, it is more efficient to utilize a single instrumentation amplifier in conjunction with a multiplexer to sequentially switch between leads. Signal filtering is more effectively performed in the

digital domain, as digital filters are not constrained by the number of cascaded stages and can be designed with greater flexibility. Compared to analog filters, digital filters provide higher adaptability, since their characteristics can be implemented in software and readily adjusted to meet specific signal processing requirements [42, 49].

6 Conclusion

The result of this review paper is to explain various models of digital filter design development on ECG but focused on IIR filters. From the results of the review paper, there are several methods of using digital filters and analysis methods. From the review of several papers, it is shown that ECG signals are more commonly found at frequencies between 0.5 Hz and 100 Hz. Several filters are used in the literature discussed in this paper, some differences between the literature and others are found. In some papers, it was found that the butterworth filter is the best and most stable filter for ECG tools. In addition to butterworth filters, there are several types of filters that can be used such as Chebyshev type 1, Chebyshev type 2, Elliptic and many more types of digital filters. In terms of product processing, product processing in the matlab application is a widely used choice because of its complete features and easy-to-use interface. There are several methods that can be done using the matlab application. The methods that are often encountered are SVR, MSE, SIR, PSD and FFT. In general, data processing systems using sample frequency still show good accuracy with low computing time. Some of the things that need to be developed in ECG research are the use of digital filters that have the best filters to compare with the analog filters that are best used and the use of internet-based alarm systems to prevent heart attacks in humans and for good diagnosis.

References

1. Abrar M, Khan WE (2018) Digital filtering techniques for power-line interference removal from ECG signals. In: 2018 Conference on information communications technology and society ICTAS 2018—Proceedings, pp 1–6. https://doi.org/10.1109/ICTAS.2018.8368750
2. Aiboud Y et al (2016) Review of ECG signal de-noising techniques. In: Proceedings of 2015 IEEE World conference on complex systems WCCS 2015. https://doi.org/10.1109/ICoCS.2015.7483313
3. Almalchy MT et al (2019) Noise removal from ECG signal based on filtering techniques. In: Proceedings—2019 22nd International conference on control systems and computer science CSCS 2019, pp 176–181. https://doi.org/10.1109/CSCS.2019.00037
4. Altay YA et al (2019) The effect of filter parameters on the accuracy of ECG signal measurement. Biomed Eng (NY) 53(3):176–180. https://doi.org/10.1007/s10527-019-09903-2
5. Ashida T et al (2017) Usefulness of synthesized 18-lead electrocardiography in the diagnosis of STEMITOC and title pages. Am J Emerg Med 35(3):448–457. https://doi.org/10.1016/j.ajem.2016.11.052
6. Aulia F Implementasi low pass filter digital IIR (infinite-impulse response) butterworth pada FPGA, 2
7. Buendía-Fuentes F et al (2012) High-bandpass filters in electrocardiography: source of error in the interpretation of the ST segment. ISRN Cardiol 2012:1–10. https://doi.org/10.5402/2012/706217

8. Burguera A (2019) Fast QRS detection and ECG compression based on signal structural analysis. IEEE J Biomed Heal Informatics 23(1):123–131. https://doi.org/10.1109/JBHI.2018.2792404
9. Butterworth S (1930) On the theory of filter amplifiers
10. Chavan MS et al (2008) FIR Equiripple digital filter for reduction of power line interference in the ECG signal, 147–150
11. Coviello JS (2015) ECG interpretation made incredibly easy! 6th edn.
12. Dong J, Jiang W (2016) Design of digital filter on ECG signal processing. In: Proceedings—5th International conference on instrumentation & measurement, computer, communication, and control IMCCC 2015, pp 1272–1275. https://doi.org/10.1109/IMCCC.2015.273
13. Faiz MMU, Kale I (2022) Removal of multiple artifacts from ECG signal using cascaded multistage adaptive noise cancellers. Array 14:100133. https://doi.org/10.1016/j.array.2022.100133
14. Gaikwad KM, Chavan MS (2014) Removal of high frequency noise from ECG signal using digital IIR butterworth filter. In: Proceedings—2014 Global conference on wireless computing and networking, GCWCN 2014, pp 121–124. https://doi.org/10.1109/GCWCN.2014.7030861
15. Gargiulo GD et al (2018) On the einthoven triangle: a critical analysis of the single rotating dipole hypothesis. Sensors (Switzerland) 18(7). https://doi.org/10.3390/s18072353
16. Gargiulo GD (2015) True unipolar ECG machine for Wilson central terminal measurements. Biomed Res Int. https://doi.org/10.1155/2015/586397
17. Handayani A (2017) Sistem Konduksi Jantung. Bul Farmatera 2(3):116. https://doi.org/10.30596/bf.v2i3.1197
18. Hardi W, Wangko S (2013) Peran Sel Nodus Sinoatrial Sebagai Pengatur Irama Jantung. J Biomed 4(3):35–41. https://doi.org/10.35790/jbm.4.3.2012.1212
19. Hary (2010) Real-time filter FIR. Pengolah Sinyal, 1–12
20. Hu X et al (2010) Reduction arithmetic for power line interference from ecg based on estimating sinusoidal parameters. In: Proceedings—2010 3rd International conference on biomedical engineering and informatics, BMEI 2010, vol 5. BMEI, pp 2089–2092. https://doi.org/10.1109/BMEI.2010.5640006
21. Ingale M et al (2020) ECG biometric authentication: a comparative analysis. IEEE Access 8:117853–117866. https://doi.org/10.1109/ACCESS.2020.3004464
22. Jagtap SK, Uplane MD (2012) The impact of digital filtering to ECG analysis: Butterworth filter application. In: Proceedings—2012 International conference on computer and communication technology, ICCICT 2012. pp 1–6. https://doi.org/10.1109/ICCICT.2012.6398145
23. Jazlan A et al (2013) A review on reduced order approximation for digital filters with complex coefficients using model reduction. In: 2013 3rd Australian control conference, AUCC 2013, pp 79–84. https://doi.org/10.1109/AUCC.2013.6697251
24. Keshavamurthy TG, Eshwarappa MN (2017) Review paper on denoising of ECG signal. In: Proceedings of 2017 Second international conference on electrical, computer and communication technologies ICECCT 2017. https://doi.org/10.1109/ICECCT.2017.8117941
25. Kora P et al (2017) ECG based atrial fibrillation detection using Sequency ordered complex Hadamard transform and hybrid firefly algorithm. Eng Sci Technol an Int J 20(3):1084–1091. https://doi.org/10.1016/j.jestch.2017.02.002
26. Lan T, Pham N (ECG) Signal using digital signal processing (DSP) with
27. Li Q et al (2009) Artificial arterial blood pressure artifact models and an evaluation of a robust blood pressure and heart rate estimator. Biomed Eng Online 8:1–15. https://doi.org/10.1186/1475-925X-8-13

28. Lidyawati L et al (2016) Implementasi filter infinite impulse response (Iir) Dengan Respon Butterworth Dan Chebyshev Menggunakan Dsk Tms320C6713. J Elektro dan Telekomun Terap 2(1):95–103. https://doi.org/10.25124/jett.v2i1.97
29. Liu H et al (2019) Detection of fetal ECG R wave from single-Lead abdominal ECG using a combination of RR time-series smoothing and template-matching approach. IEEE Access 7:66633–66643. https://doi.org/10.1109/ACCESS.2019.2917826
30. Luo S, Johnston P (2010) A review of electrocardiogram filtering. J Electrocardiol 43(6):486–496. https://doi.org/10.1016/j.jelectrocard.2010.07.007
31. Miao F et al (2015) A wearable context-aware ECG monitoring system integrated with built-in kinematic sensors of the smartphone. Sensors (Switzerland) 15(5):11465–11484. https://doi.org/10.3390/s150511465
32. Mihov GS, Badarov DH (2017) Testing of digital filters for power-line interference removal from ECG signals. In: 2017 26th International scientific conference electronics 2017—Proceedings, 1–6 Jan 2017. https://doi.org/10.1109/ET.2017.8124368
33. Mittal A, Rege A (2016) Design of digital FIR filter implemented with window techniques for reduction of power line interference from ECG signal. In: IEEE International conference on communications, control, and computing, IC4 2015, pp 2–5. https://doi.org/10.1109/IC4.2015.7375535
34. Nugroho ES (2007) Pengenalan Pola Sinyal Elektrokardiograf (EKG) Dengan Jaringan Syaraf Tiruan Backpropagation Untuk Diognosa Kelainan Jantung Manusia (Perangkat Lunak). Appl Eng Semin 28–34
35. Parola F, García-Niebla J (2017) Use of high-pass and low-pass electrocardiographic filters in an international Cardiological community and possible clinical effects. Adv J Vasc Med 2(1):34–38
36. Preethi D, Valarmathi RS (2018) 2018 International conference on advances in communication and computing technology, ICACCT 2018, pp 265–268
37. Quiroz-Juarez MA et al (2018) Cardiac conduction model for generating 12 Lead ECG signals with realistic heart rate dynamics. IEEE Trans Nanobioscience 17(4):525–532. https://doi.org/10.1109/TNB.2018.2870331
38. Qureshi R et al (2017) Multistage adaptive filter for ECG signal processing. In: Proceedings of 2017 International conference on computers, communications, and systems, C-CODE 2017, pp 363–368. https://doi.org/10.1109/C-CODE.2017.7918958
39. Rachman RA et al (2021) Development of a low-cost and Effisient ECG devices with IIR digital filter design. Indones J Electron Electromed Eng Med Informat 3(1):21–28. https://doi.org/10.35882/ijeeemi.v3i1.4
40. Raita-aho T et al (1994) A digital filter Chip for ECG signal processing. IEEE Trans Instrum Meas 43(4):644–649. https://doi.org/10.1109/19.310181
41. Redmond SJ et al (2008) ECG quality measures in telecare monitoring. In: Proceedings of the 30th annual international conference of the IEEE engineering in medicine and biology society EMBS'08—Personalized healthcare through technology, pp 2869–2872. https://doi.org/10.1109/iembs.2008.4649801
42. Ribatti D (2009) William Harvey and the discovery of the circulation of the blood. J Angiogenes Res 1(1):3. https://doi.org/10.1186/2040-2384-1-3
43. Satija U et al (2018) A review of signal processing techniques for electrocardiogram signal quality assessment. IEEE Rev Biomed Eng 11, c:36–52. https://doi.org/10.1109/RBME.2018.2810957
44. Saxena C et al (2018) Denoising of ECG signals using FIR & IIR filter: a performance analysis. Int J Eng Technol 7(4):1–5. https://doi.org/10.14419/ijet.v7i4.12.20982
45. Selvaraj S et al (2021) Filtering the ECG signal towards heart attack detection using motion artifact removal technique. In: Proceedings of the third international conference on intelligent

communication technologies and virtual mobile networks, ICICV 2021, pp 185–188. https://doi.org/10.1109/ICICV50876.2021.9388515
46. Singh N et al (2013) Design of digital IIR filter for noise reduction in ECG signal. In: Proceedings—Fifth international conference on computing and communication networks, CICN 2013, pp 171–176. https://doi.org/10.1109/CICN.2013.45
47. Sornmo L, Laguna P ((206)AD) Electrocardiogram (ECG) signal processing. Biomed Eng 28(1)
48. Spanò E et al (2016) Low-power wearable ECG monitoring system for multiple-patient remote monitoring. IEEE Sens J 16(13):5452–5462. https://doi.org/10.1109/JSEN.2016.2564995
49. Sudjadi I et al (2011) Pengenalan Pola Sinyal Elektrokardiograf (EKG) dengan Jaringan Syaraf Tiruan Backpropagation untuk Diagnosa Kelainan Jantung Manusia
50. Tomasic I, Trobec R (2014) Electrocardiographic systems with reduced numbers of leads-synthesis of the 12-Lead ECG. IEEE Rev Biomed Eng 7:126–142. https://doi.org/10.1109/RBME.2013.2264282
51. Tompkins WJ (2000) Biomedical digital signal processing book. https://doi.org/10.1007/SpringerReference_12454
52. Trigo JD et al (2012) A review on digital ECG formats and the relationships between them. IEEE Trans Inf Technol Biomed 16(3):432–444. https://doi.org/10.1109/TITB.2011.2176955
53. Tseng CC, Lee SL (2019) Closed-form design of digital FIR notch filter using chebyshev polynomial. In: 2019 IEEE 8th Global conference on consumer electronics, GCCE 2019, pp 287–288. https://doi.org/10.1109/GCCE46687.2019.9015306
54. Zhang X, Jiang S (2021) Application of Fourier transform and butterworth filter in signal denoising. In: 2021 IEEE 6th International conference on intelligent computing and signal processing, ICSP 2021, pp 1277–1281. https://doi.org/10.1109/ICSP51882.2021.9408933

Author Index

A
Abdullayev, Vugar 521
Adi, Rizky 474
Ahmad, Umar Ali 474
Ahniar, Nur Hasanah 459
Ain, Khusnul 146
Alfisyah, Rima Firda 146
Ali, Kharudin 336
Alsari, Reno Adi Surya 319
Ama, Fadli 146
Amalia, Qoriina Dwi 81
Amitasari, Nisa R. 53
Amrustian, Muhammad Afrizal 109
Andrianto, Siddiq 459
Ani, Adi Izhar Che 378
Anifah, Lilik 279, 291
Arias, José Luis Ramírez 219
Arifah, Nur Latifatul 184
Aris, Raja Siti Nur Adiimah Raja 336
Asmaria, Talitha 53
As-Sidqi, Luthfi 305
Athavale, Vijay Anant 507
Ayati, Nonik S. 53
Aziz, Sulthon Abdul 429
Azizah, Wafiq 27

B
Balqis, Fahriza 378
Basari 429

C
Caesarendra, Wahyu 494
Chai, Rifai 146
Cuangdinata, William Axel 92

D
Danuartha, I Dewa Putu Arvian 172
Delimayanti, Mera Kartika 474
Dewa Gede Hari Wisana, I. 361
Dewiningrum, Riqqah 361

E
Erfianto, Bayu 81

F
Fa'altin, Tetrik 172
Fakhrurreza, Muhammad 446
Fauziah, Syifa 485
Fredson, Kent Alber 92

G
Gafur, Fira Angelia 71
Ghoni, Ruzlaini 1, 393

H
Hadi, Mukhammad Nur 378
Hapsari, Rinci Kembang 202
Haryanto 279, 291
Hermawan, Norma 250
Hidayati, Rinda Nur 459
Hikmah, Nada Fitrieyatul 235, 250
Huynh, Phuoc-Hai 121, 160

I
I'Tikap, Muhammad 485
Ibrahim, Mohd Tarmizi 393
Irianto, Bambang Guruh 319
Ismail, Rifky 133
Istiqomah 81, 305
Istiqomah, Anisa Nur 446
Ittaqillah, Sayyidul Istighfar 146

J
Joharwan, Johanes Wawan 1

K
Khayr, Rashad Abul 418
Kholiq, Abd. 319
Kurniatie, Menik Dwi 53
Kusumawardana, Arya 378, 402
Kuswanto, Dwi 291

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2026
Triwiyanto et al. (Eds.): icebehi 2024, LNEE 1611, pp. 541–543, 2026.
https://doi.org/10.1007/978-981-95-9233-3

L
Lestari, Gita Rindang 459
Lestari, Tri Riana 485
Luthfiyah, Sari 507

M
Maghfiroh, Anita Miftahul 172, 319, 459, 494
Maharani, Puan Ayu 12
Mayasari, Dita A. 53
Meiliana 92
Mohamed, Abdul Halim Ikram 393
Mule, Nilakshi Maruti 263
Musvika, Syevana Dita 319

N
Nabila, Noor Anisa 40
Nafianto, Muhammad Abyan 172
Nazaruddin, Yul Yunazwin 418
Novriansyah, Robin 133
Nugraha, Priyambada Cahya 172, 361
Nugraha, Rajendra Andita 446
Nugroho, Wasis 1, 336
Nur, Sulistia 133
Nurhayati 279
Nurromadhona, Salma Syahidah 172

O
Octavia, Atin Asna 146
Oktavianti, Euis 474

P
Pakaya, Andini Vira Salsabilla Zulkarnain 250
Pangaribuan, Porman 418
Passa, Emmanuella Aurelia Rachel 202
Patil, Dipti Durgesh 263
Pharmawati, Kancitra 305
Prabowo, Riefky 27
Prawibowo, Hartanto 133
Purnama, Priska Arindya 71
Purnamasari, Dian Neipa 291
Purnomo, Agung 12, 27, 40
Purwanti, Endah 202
Putra, Alfian Pramudita 202
Putra, Raden Achmad Candra 485
Putri, Farika Tono 133
Putu Alit Pawana, I. 507

R
Raharjo, Emanuel Budi 1
Rahma, Osmalina Nur 146
Rahmawati, Triana 521
Ramadhan, Ichsan 1
Ratri, Clara Tyas 146
Rifai, Damhuji 1, 336
Rizal, Achmad 81
Rizqullah, Muhammad Akmal Falih 305
Rogo, Danu 361
Romdlony, Muhammad Zakiyullah 418
Rosas, Daniel Botero 219
Rusimamto, Puput Wanarti 279

S
Sabrina, Febiola 402
Saffinee, Siti Nurshafina Zaharah 336
Samatha, Syifa Candiki 202
Samsuri, Fahmi 336
Saputra, Muhammad Ikhwani 109
Sardjono, Tri Arief 235, 250
Satou, Kenji 474
Setyoningsih, Endang Dian 521
Shanti, Mardhatillah 12
Shavin, Mohamad 235
Solórzano, Giovanny Andrés Piedrahita 219
Subyakto 485
Suratman, Fiky Y. 305
Susanto, Husein Arif 1
Syaifudin 184

T
Tamba, Tua Agustinus 418
Tauviqirrahman, Mohammad 133
Triwardono, Joko 53
Triwiyanto 494, 507, 521

U
Utomo, Bedjo 507
Utomo, Muhammad S. 53

W
Wakidi, Levana Forra 361, 507
Wardhani, Eka 305
Warsuta, Bambang 474
Widjaya, Mikhail Sannata 305
Wisana, I Dewa Gede Hari 172

Y
Yogatama, Andre Dimas 402
Yulianto, Endro 184, 494

Z
Zakiyyah, Alfi Yusrotis 92
Zen, Bita Parga 109
Zuhrie, Muhamad Syarifuddien 279