Lecture Notes in Networks and Systems 445

Raghvendra Kumar
Prasant Kumar Pattnaik
João Manuel R. S. Tavares *Editors*

Next Generation of Internet of Things

Proceedings of ICNGIoT 2022

Springer

Lecture Notes in Networks and Systems

Volume 445

The series “Lecture Notes in Networks and Systems” publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Raghvendra Kumar · Prasant Kumar Pattnaik ·
João Manuel R. S. Tavares
Editors

Next Generation of Internet of Things

Proceedings of ICNGIoT 2022

Springer

Editors
Raghvendra Kumar
GIET University
Gunupur, India

Prasant Kumar Pattnaik
KIIT University
Bhubaneswar, Odisha, India

João Manuel R. S. Tavares
Faculdade de Engenharia da
Universidade do Porto
Porto, Portugal

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-19-1411-9 ISBN 978-981-19-1412-6 (eBook)
https://doi.org/10.1007/978-981-19-1412-6

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

Preface

This book collects a high-quality, peer-reviewed papers presented at the second international conference on next generation of Internet of Things, organized by the department of computer science and engineering, GIET University, Gunupur-765022, India, during February 03–04, 2022. The book is organized with 60 chapters and discusses advanced and multi-disciplinary research regarding the next generation of IoT and its applications. It focuses on innovative technologies and application in different areas of artificial intelligence, machine learning, deep learning, data analytics, data science, computation technologies, cloud and mobile computing, SVM, networking, healthcare, bioinformatics, and biomedical, etc. Further, the book also addresses the deployment of IoT in computational technologies and information transfer approaches and solutions in various disciplines of science, engineering, and technology.

Gunupur, India — Raghvendra Kumar
Bhubaneswar, India — Prasant Kumar Pattnaik
Porto, Portugal — João Manuel R. S. Tavares

Contents

Editors and Contributors

About the Editors

Dr. Raghvendra Kumar is working as Associate Professor in Computer Science and Engineering Department at GIET University, India. He received B. Tech., M.Tech., and Ph.D. in Computer Science and Engineering, India, and Postdoc Fellow from Institute of Information Technology, Virtual Reality and Multimedia, Vietnam. He serves as Series Editor *Internet of Everything (IOE): Security and Privacy Paradigm*, *Green Engineering and Technology: Concepts and Applications*, published by CRC press, Taylor & Francis Group, USA, and *Bio-Medical Engineering: Techniques and Applications*, published by Apple Academic Press, CRC Press, Taylor & Francis Group, USA. He also serves as Acquisition Editor for *Computer Science* by Apple Academic Press, CRC Press, Taylor & Francis Group, USA. He has published number of research papers in international journals (SCI/SCIE/ESCI/Scopus) and conferences including IEEE and Springer as well as serve as Organizing Chair (RICE-2019, 2020), Volume Editor (RICE-2018), Keynote Speaker, Session Chair, Co-chair, Publicity Chair, Publication Chair, Advisory Board Member, Technical Program Committee Member in many international and national conferences and serve as Guest Editor in many special issues from reputed journals (Indexed By: Scopus, ESCI, SCI). He also published 13 chapters in edited book published by IGI Global, Springer, and Elsevier. His researches areas are computer networks, data mining, cloud computing and secure multiparty computations, theory of computer science, and design of algorithms. He authored and edited 23 computer science books in field of Internet of Things, data mining, biomedical engineering, big data, robotics, and IGI Global Publication, USA, IOS Press Netherland, Springer, Elsevier, CRC Press, USA.

Prasant Kumar Pattnaik, Ph.D. (Computer Science), Fellow IETE, Senior Member IEEE is Professor at the School of Computer Engineering, KIIT Deemed University, Bhubaneswar. He has more than a decade of teaching and research experience and awarded half dozen of Ph.D. Dr. Pattnaik has published numbers of research papers in peer-reviewed international journals and conferences and filed many patents. He also edited book volumes in Springer and IGI Global Publication. His areas of interest include mobile computing, cloud computing, cyber security, intelligent systems, and brain–computer interface. He is one of the associate editors of *Journal of Intelligent and Fuzzy Systems*, IOS Press, and *Intelligent Systems* Book Series Editor of CRC Press, Taylor Francis Group.

João Manuel R. S. Tavares graduated in Mechanical Engineering at the Universidade do Porto, Portugal, in 1992. He also earned his M.Sc. degree and Ph.D. degree in Electrical and Computer Engineering from the Universidade do Porto in 1995 and 2001 and attained his habilitation in Mechanical Engineering in 2015. He is Senior Researcher at the Instituto de Ciência e Inovação em Engenharia Mecânica e Engenharia Industrial (INEGI) and Associate Professor at the Department of Mechanical Engineering (DEMec) of the Faculdade de Engenharia daUniversidade do Porto (FEUP). João Tavares is Co-editor of more than 60 books and Co-author of more than 50 chapters, 650 articles in international and national journals and conferences, and 3 international and 3 national patents. He has been Committee Member of several international and national journals and conferences, is Co-founder and Co-editor of the book series "Lecture Notes in Computational Vision and Biomechanics" published by Springer, Founder and Editor-in-Chief of the journal *Computer Methods in Biomechanics* and *Biomedical Engineering: Imaging and Visualization* published by Taylor & Francis, Editor-in-Chief of the journal *Computer Methods in Biomechanics and Biomedical Engineering* published by Taylor & Francis, and Co-founder and Co-chair of the international conference series: CompIMAGE, ECCOMAS VipIMAGE, ICCEBS, and BioDental. Additionally, he has been (co-)Supervisor of several M.Sc. and Ph.D. thesis and Supervisor of several postdoc projects and has participated in many scientific projects both as Researcher and as Scientific Coordinator. His main research areas include computational vision, medical imaging, computational mechanics, scientific visualization, human–computer interaction, and new product development.

Contributors

Abbas Khalaf Haitham University of Anbar, Ramadi, Iraq

Abbas Zainab Adnan University of Technology, Baghdad, Iraq

Abdulbaqi Azmi Shawkat Department of Computer Science, College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq

Abdulrasool Hasan Husam Al-Esraa University College, Air Conditioning and Refrigeration Department, Baghdad, Iraq

Abdulrazzaq Musaab Riyadh Department of Computer Science, Mustansiriyah University, Baghdad, Iraq

Agarwal Parul Jamia Hamdard, New Delhi, India

Al-Anbary Karrar M. College of Engineering, Electrical Engineering Department, University of Babylon, Babel, Iraq

Al-Itbi Ali Sami Informatics Department, Technical College of Management, Middle Technical University, Baghdad, Iraq

Al-Janabi Rana Jumaa Sarih Computer Science Department, University of Alqadisiyah, qadisiyah, Iraq

Al-Obaidi Fatin E. M. Department of Physics, College of Science, Mustansiriyah University, Baghdad, Iraq

Al-Shimiry Intisar F. H. Department of Physics, College of Science, Mustansiriyah University, Baghdad, Iraq

Al-Zuky Ali A. D. Department of Physics, College of Science, Mustansiriyah University, Baghdad, Iraq

Alanezi Mafaz Department of Computer Science, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq

Ali Sarah Fouad Department of Computer Science, Mustansiriyah University, Baghdad, Iraq

Alqassab Asmaa Department of Computer Science, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq

Anh Tran Tuan Thu Dau Mot University, Thu Dau Mot, Vietnam

Anil Kumar Dodala GIET University, Gunupur, Odisha, India

Ansaf Alla Hussein Omran Kut Technical Institute, Middle Technical University, Baghdad, Iraq

Anumotu Sitara Department of Computer Science and Engineering, Manav Rachna International Institute of Research and Studies, Faridabad, Haryana, India

Apat Shraban Kumar Department of CSE, GIET University, Gunupur, Odisha, India

Arnoos Salam Muhsin Informatics Department, Technical College of Management, Middle Technical University, Baghdad, Iraq

Arpana D. The Oxford College of Business Management, Bangalore, India

Arya Rajeev Wireless Sensor Networks Lab, Department of Electronics and Communication Engineering, National Institute of Technology Patna, Bihar, India

Ashraf Faizan GIET University, Gunupur, India

Ayem Gabriel Terna American University of Nigeria, Yola, Nigeria

Aziz Ameen Kayser Al-Esraa University College, Air Conditioning and Refrigeration Department, Baghdad, Iraq

Babu Sudheer GIET University, Gunupur, Odisha, India

Badyal Bhanvi SCIT, Manipal University Jaipur, Jaipur, India

Balhara Amit Department of Computer Science and Engineering, Manav Rachna International Institute of Research and Studies, Faridabad, Haryana, India

Bondugula Rohit Kumar School of Computer and Information Sciences, University of Hyderabad, Hyderabad, India

Chawla Pronika Department of Computer Science and Engineering, Manav Rachna International Institute of Research and Studies, Faridabad, Haryana, India

Chi Nguyen Van Thai Nguyen University of Technology, Thai Nguyen, Vietnam

Dansana Debabrata Rajendra University, Balangir, India

Das Kalyan Sambalpur University Institute of Information Technology, Sambalpur, India

Das Rajendra Kumar Department of CSE, DRIEMS, Cuttack, India;
Department of ENTC, DRIEMS, Cuttack, India

Das Satya Narayan Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India

Dash Sachikanta Department of CSE, DRIEMS, Cuttack, India

Digaari Hrishabh SCIT, Manipal University Jaipur, Jaipur, India

Duc Ngo Minh Thai Nguyen University of Technology, Thai Nguyen, Vietnam

Duong Pham Dinh Hanoi University of Science and Technology, Hanoi, Vietnam

Enaw Hayder H. College of Engineering, Electrical Engineering Department, University of Babylon, Babel, Iraq

Gaata Methaq Talib Department of Computer Science, Mustansiriyah University, Baghdad, Iraq

Gupta Kamaldeep Faculty of Computing and Information Technology, Usha Martin University, Ranchi, India

Gupta Rishi SCIT, Manipal University Jaipur, Rajasthan, India

Gupta Shaurya Department of Computer Science and Engineering, UPES, Dehradoon, India

Hadi Ahmed Qassim University of Technology, Baghdad, Iraq

Hamidon Mohd Nizar Electrical and Electronic Engineering Department and Deputy Director of Institute of Advanced Technology at University Putra Malaysia (UPM), Putra, Malaysia

Hammadi Khaleel J. Department of Computer Techniques Engineering, Al-Esraa University College, Baghdad, Iraq

Handa Disha University Institute of Computing, Chandigarh University, Mohali, India

Hasan Hiba A. Computer Techniques Engineering Department, Al-Turath University College, Baghdad, Mansour, Iraq

Hatam Ali A. Al-Esraa University College, Air Conditioning and Refrigeration Department, Baghdad, Iraq

Hilal Zahraa Mohammed University of Technology, Baghdad, Iraq

Ho-Dac Hung Thu Dau Mot University, Thu Dau Mot, Vietnam

Hussein Hafedh Hameed Karbala Education Directorate, Karbala, Iraq

Hussein Karim Q. Computer Science Department, University of Mustansiriyah, Baghdad, Iraq

Ibrahim Amer A. Department of Computer Techniques Engineering, Al-Esraa University College, Baghdad, Iraq

Idrees Sheikh Mohammad Norwegian University of Science and Technology, Trondheim, Norway

Ilayaraja M. Department of Computer Applications, Kalasalingam Academy of Research and Education, Virudhunagar, Tamil Nadu, India

Jain Tarun SCIT, Manipal University Jaipur, Jaipur, Rajasthan, India

Jawad Ghali Rana Computer Science Department, University of Mustansiriyah, Baghdad, Iraq

Jena U. K. Department of Computer Science and Engineering, VSSUT, Burla, Odisha, India

Jha Kushagra Department of Computer Science and Engineering, Manav Rachna International Institute of Research and Studies, Faridabad, Haryana, India

Kabat M. R. Department of Computer Science and Engineering, VSSUT, Burla, Odisha, India

Kadhum Hayder Hussein College of Engineering, Electrical Engineering Department, University of Babylon, Babel, Iraq

Kamil Flayyih Noora University of Babylon, Babil/Hilla, Iraq

Kaur Prabhsharan Chandigarh University, Mohali, India

Khaleel Hasan H. Department of Medical Devices Techniques Engineering, Al-Esraa University College, Baghdad, Iraq

Kiran Kumar Reddy P. MRIET-JNTU, Hyderabad, India

Kishore Mishra Brojo Department of Artificial Intelligence and Data Science, Vardhaman College of Engineering (Autonomous), Hyderabad, India

Kuanar Sanjay Kumar School of Engineering and Technology, Department of Computer Science and Engineering (CSE), GIET University, Gunupur, Odisha, India

Kumar Brijesh Manav Rachna International Institute of Research and Studies, Faridabad, India

Kumar Das Pradipta Department of Information Technology, VSSUT, Burla, Odisha, India

Kumar Lov Department of Computer Science and Information Systems, BITS-Pilani, Hyderabad, India

Kumar Pankaj Chandigarh University, Mohali, India

Len Vo Van Thu Dau Mot University, Thu Dau Mot, Vietnam

Linh Bui Thi Hai Thai Nguyen University of Technology, Thai Nguyen, Vietnam

Madaan Geetika University Centre for Research and Development (UCRD), Chandigarh University, Ludhiana, Punjab, India

Mahmood Sawsan Dheyaa College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq

Mekala Bhaskar GIET University, Gunupur, Odisha, India

Mishra Jyotirmaya Department of CSE, School of Engineering & Technology, GIET University, Gunupur, Odisha, India

Mittal Shilpi Chandigarh University, Mohali, Punjab, India

Mohamed Shili Innov'COM at Sup'Com Laboratory, University of Carthage, Carthage, Tunisia

Mohammed Enas Ali University of Kerbela, Karbala, Iraq

Mohammed Nibras Talib University of Kerbela, Karbala, Iraq

Mohapatra Pradyumna Kumar Research Scholar, Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India;
Vedang Institute of Technology, Bhubaneswar, Odisha, India

Mohd Maryam Biti Department of Electrical and Electronic Engineering, Universiti Putra, Putra, Malaysia

Mohy Raheem J. Al-Esraa University College, Air Conditioning and Refrigeration Department, Baghdad, Iraq

Mona Jammel University of Kufa Iraq, Najaf, Iraq

Murdas Ibrahim A. University of Babylon, Babil/Hilla, Iraq

Najim Saif Al-din M. College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq

Nam Dao Phuong Hanoi University of Science and Technology, Hanoi, Vietnam

Nejrs Salwa Mohammed Directorate of Private University Education, Ministry of Higher Education, Baghdad, Iraq

Nghia Nguyen Trung Hanoi University of Science and Technology, Hanoi, Vietnam

Obaid Ahmed J. Faculty of Computer Science and Mathematics, University of Kufa, Kufa, Iraq

Omran Eman Hatem Computer Science Department, University of Alqadisiyah, qadisiyah, Iraq

Padhy Neelamadhab Department of CSE, School of Engineering and Technology, GIET University, Gunupur, Odisha, India

Padhy Sasmita Department of CSE, KL University, Vaddeswaram, India

Panda Mrutyunjaya Department of Computer Science and Applications, Utkal University, Bhubaneswar, Odisha, India

Panda Rabinarayan Department of CSE, GIET University, Gunupur, India

Pandey Digvijay Department of Technical Education, IET, Dr. A.P.J. Abdul Kalam Technical University, Lucknow, Uttar Pradesh, India

Panigrahi Rasmita School of Engineering and Technology, Department of Computer Science and Engineering (CSE), GIET University, Gunupur, Odisha, India

Patnaik Archana School of Engineering and Technology, Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India

Patro Sibo Prasad School of Engineering and Technology, Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India

Prasad Patnaik Prabhu School of Engineering & Technology, Department of CSE, GIET University, Gunupur, Odisha, India

Prasad Sanaboina S. V. S. V. GIET University, Gunupur, Odisha, India

Sharma Kanta Prasad University Institute of Computing Chandigarh University Mohall, Mohali, Punjab, India

Prasad Vivek Kumar Nirma University, Ahmedabad, India

Quang Nguyen Hong Thai Nguyen University of Technology, Thai Nguyen, Vietnam

Rai Kajal University Institute of Computing, Chandigarh University, Mohali, India

Rajalakshmi N. R. Computer Science & Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India

Rajiv K. GRIET-JNTUH, Hyderabad, India

Ranjan Kabat Manas Department of CSE, VSSUT, Burla, Odisha, India

Rout Saroja Kumar Department of Information Technology, Vardhaman College of Engineering (Autonomous), Hyderabad, Telangana, India

Routray Nalinikanta Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India

Roy Sharmistha Faculty of Computing and Information Technology, Usha Martin University, Ranchi, India

Sadiq Mohamed Abdul Karim Department of Information Technology, University of Technology and Applied Sciences, Suhar, Sultanate of Oman

Sadiq Mohammed Dhi-Qar Investment Commission, Dhi-Qar, Iraq

Sah Rahul Deo Dr. Shyama Prasad Mukherjee University, Ranchi, India

Sahan Ali Mohammed Informatics Department, Technical College of Management, Middle Technical University, Baghdad, Iraq

Sahoo H. K. Department of Electronic and Telecomunication Engineering, VSSUT, Burla, Odisha, India

Sahoo Subhashree Centurion University of Technology and Management, Bhubaneswar, India

Sahu Bandita Department of CSE, GIET University, Gunupur, Odisha, India

Sahu Bibhuprasad Department of Information Technology, Vardhaman College of Engineering (Autonomous), Hyderabad, Telangana, India

Salahaldeen Alneamy Jamal Department of Software, College of Information Technology, Ninevah University, Mosul, Iraq;
Department of Software, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq

Samal S. Department of Electronic and Telecomunication Engineering, VSSUT, Burla, Odisha, India

Sankar S. Department of CSE, Sona College of Technology, Salem, India

Senthilkumar Sudha Department of Information Technology, University of Technology and Applied Sciences, Suhar, Sultanate of Oman

Sethom Kaouthar Innov'COM at Sup'Com Laboratory, University of Carthage, Carthage, Tunisia

Shanmugam Thirumurugan Department of Information Technology, University of Technology and Applied Sciences, Suhar, Sultanate of Oman

Shanmugarajeshwari V. Department of Computer Applications, Kalasalingam Academy of Research and Education, Virudhunagar, Tamil Nadu, India

Sharma Isha Chandigarh University, Mohali, India

Sharma Rohit Wireless Sensor Networks Lab, Department of Electronics and Communication Engineering, National Institute of Technology Patna, Bihar, India

Sheenam Chandigarh University, Mohali, India

Sherza Jenan S. Al-Esraa University College, Air Conditioning and Refrigeration Department, Baghdad, Iraq

Shrotriya Anita SCIT, Manipal University Jaipur, Rajasthan, India

Singh Jarnail University Institute of Computing, Chandigarh University, Mohali, Punjab, India

Singh Sanjeet Chandigarh University, Ludhiana, Punjab, India

Sinha Aditya SCIT, Manipal University Jaipur, Rajasthan, India

Siva Krishna Kotha GIET University, Gunupur, Odisha, India

Sivangi Kaushik Bhargav School of Computer and Information Sciences, University of Hyderabad, Hyderabad, India

Srujan Raju K. Department of Computer Science and Engineering, CMR Technical Campus, JNTUH, Hyderabad, Telangana, India;
Computer Science & Engineering, School of Engineering & Technology, GIET University, Gunupur, Odisha, India

Sulaiman Nasri Bin Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, Putra, Malaysia

Sundara Kumar M. R. Department of CSE, Sona College of Technology, Salem, India

Swapna H. R. School of Commerce- PG Studies, JAIN (Deemed-to-Be University), Bangalore, India

Thakkar Abhi K. Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Thandekkattu Salu George American University of Nigeria, Yola, Nigeria

Thanh Hung Bui Artificial Intelligence Laboratory, Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam

Thanoon ALkahla Lubna Department of Computer Science, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq;

Department of Software, College of Information Technology, Ninevah University, Mosul, Iraq

Thuy Nguyen Vinh Thai Nguyen University of Technology, Thai Nguyen, Vietnam

Tiwari Pradeep Kumar SCIT, Manipal University Jaipur, Rajasthan, India

Udgata Siba K. School of Computer and Information Sciences, University of Hyderabad, Hyderabad, India

Ukani Vijay Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Vajjhala Narasimha Rao University of New York Tirana, Tirana, Albania

Van Trong Dang Hanoi University of Science and Technology, Hanoi, Vietnam

Vashisht Manisha Manav Rachna International Institute of Research and Studies, Faridabad, India

Walia Kirti University Institute of Computing Chandigarh University Mohall, Mohali, India

IoT-Assisted Crop Monitoring Using Machine Learning Algorithms for Smart Farming

Shraban Kumar Apat, Jyotirmaya Mishra, K. Srujan Raju, and Neelamadhab Padhy

Abstract Agriculture expansion is critical to the economic prosperity of any country. Agriculture employs more than 60% of the Indian population, either directly or indirectly. Nowadays, monitoring the crop is the challenging task in the world. In this article, data has been collected from various sensors to propose an IoT-assisted hybrid machine learning approach for obtaining an effective crop monitoring system. Crop monitoring system here means predicting as well as detecting diseases of crops. This study is about leveraging existing data and applying regression analysis, SVM, and decision tree to predict crop diseases in diverse crops such as rice, ragi, gram, potato, and onion. Among the applied methods, SVM outperforms regression, DT methods. The training and testing accuracy of Gram has 96.29% and 95.67%, respectively.

Keywords SVM · NB · Crop monitoring · IoT-assisted sensors

1 Introduction

Agriculture research is rapidly expanding as a result of technological advancements. It provides numerous academics with the potential to address upcoming agricultural difficulties. Humidity, temperature, rainfall, wind speed, light intensity, and soil pH are all important environmental elements in precision agriculture. Smart farming is

S. Kumar Apat (✉) · J. Mishra · K. Srujan Raju · N. Padhy
Computer Science & Engineering, School of Engineering & Technology, GIET University, Gunupur 765022, Odisha, India
e-mail: shraban.apat@giet.edu

J. Mishra
e-mail: jyoti@giet.edu

N. Padhy
e-mail: dr.neelamadhab@giet.edu

K. Srujan Raju
Department of Computer Science and Engineering, CMR Technical Campus, Hyderabad, Telangana, India

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_1

vital for dealing with a variety of difficulties in the agriculture area, such as production, early disease detection, environmental effect, and climatic circumstances. Crops that are healthy maximize productivity; otherwise, farmers will lose a lot of money. In whole-crop production, it is critical to identify the pest. Pest control sensors keep an eye on the pest population as well as the surrounding environment. Farmers can follow the full history of pest attacks by using IoT-enabled pest control devices. The appropriate and proper quantity of pesticides must be automatically sprayed in the field based on this data. Wireless sensor networks track and detect insect populations, automatically monitor chemical delivery systems, and activate and maintain crop health. The comprehensive study of machine learning principles, algorithms, and methodologies is described in this chapter. This article provides an overview of the suggested disease detection and prevention strategy in agriculture. In whole-crop production, identifying the pest is critical. Pest control sensors keep track of both the pest population and the environment. Farmers can keep a complete record of pest attacks by implementing IoT-enabled pest control sensors. Based on this information, the required and appropriate quantity of pesticides must be automatically sprayed into the field. Wireless sensor networks can detect and monitor pest populations, automatically monitor chemical delivery systems, and activate and maintain crop health. This study investigate ML algorithm for obtaining an effective crop monitoring system. Crop monitoring systems are used to both predict and detect crop diseases. This chapter also covers crop management operations such as crop yield prediction and disease detection.

2 Literature Survey

Yang et al. [1] Machine learning is being used to uncover plant resistance genes and classify plant diseases. To achieve the best classification accuracy, a careful selection of preprocessing data approaches and machine learning technologies was applied. To forecast essential plant resistance genes, more machine learning-based methods are required. To identify bacterial pathogens with high prediction precisions, ML methods such as SVM, Bayesian classifier, and RF were used (23–25). Deep convolution neural networks (CNN), the most recent generation of machine learning technologies, were used. Jawade et al [2] The field sensors gathered real-time weather data in order to anticipate disease in real time. The disease outbreak probability was calculated using the RF Regression model, which was trained on historical meteorological data. In terms of disease predicting, the model produced fairly accurate findings. The random forest algorithm has been shown to be fairly accurate in forecasting the possibility of a thrips attack. This technique has aided farmers in taking preventative steps, resulting in increased farm output and a reduction in the use of chemical pesticides on crops. Akulwar et al. [3] presents an example of how a recommender system is used in agriculture to detect and predict disease. Image preprocessing comprises noise removal and region of interest extraction. After preprocessing, feature extraction will be performed. Following the extraction of features, a ML algorithm is used, and the outcomes are anticipated. To make the best recommendations, the recommender system is used.

Ramesh et al. [4] crop selection methods are used to boost crop productivity. I used two alternative approaches. To assess datasets, the first is the Naive Bayes approach, and the second is the K-Nearest Neighbor method. Bhawana et al. [5] machine learning is being used in wheat crop production. The feature extraction of that digital image is done using digital image processing techniques. With age, the green pigment in wheat crops decreases. Supervised machine learning is used to classify crop growth stages. Technique SVM. Bondre and colleagues [6] Mahagaonkar. ML algorithm such as SVM and RF were applied to agricultural data, and a fertilizer recommendation for a specific crop was made. The aim of this research is to analyze one crop forecasting model which can be used in the latter stage. Crop yield data was gathered over the last five years from a variety of sources. The proposed work is divided into two stages: 1. Soil classification 2. Crop Yield Recommendation for Fertilizer Prediction Weather, temperature, humidity, and atmospheric pressure are all factors to consider. Third-party applications are used to display press information.

Bhawana et al. [5] the feature extraction of that digital image is done using digital image processing techniques. With age, the green pigment in wheat crops decreases. As a result, the wheat crop maturities were investigated. They determined the crop's growth stages and classified it using the supervised machine learning approach SVM. Optimizations along with fine-tuned choices of the ML techniques were used. Kumar et al. [7] the best crop yield model was created using a random forest and a decision tree. The suggested approach focuses on crop kind, yield, and weather predictions. The information in the collection is based on agricultural statistics. This dataset serves as a testing ground. After the data has been processed, it is broken into two subtypes: i.e., testing and training. Temperature, rainfall, humidity, and ph are among the variables in the dataset [9]. The datasets were referred from the Kaggle, which contains 3101 instances or data that were derived from historical data with the following parameters like humidity, rainfall Temp, ph, and crop.

Archana et al. [8] used Biosensing methods and DNA sequencing technologies to detect fungal crop diseases. Pathogenic fungus causes diseases like coils, scab, gall, leaf spot, root rot, and mildew, by making use of third-party applications. Concluded Systemic foliar infections are a major source of production and crop losses, as well as quality of crop degradation Kumar et al. [9]. Techniques such as SVM were used to recommend the suitable crop to the farmer considering environmental condition and crop disease prediction, and to suggest ways to protect them from illness Fenu et al. [10]. Specific methodologies and procedures, along with performance parameters and preprocessing techniques, were utilized. The study's findings show that this approach is still in its early stages, with significant obstacles to overcome. Data were collected from different databases such as IEEE Xplore, Science Direct, MDPI, Hindawi, and Web of Science. Over the last ten years, this study looked at classified forecasting models for crop disease (2010–2020). Aggarwal et al. [11] used Traditional agricultural procedures are being revolutionized by modern agricultural techniques that use IoT and AI, and farming is becoming a profitable venture as well. The study explains how AI and IoT might boost agricultural output by presenting several recommendations and implementation strategies. Singh et al. [11] this research leads to the creation of the Crop, a smart phone app based on a trained illness prediction model. Farmers can use the app to take crop photos and analyze the presence or absence of disease. ResNet 50, AlexNet, and ResNet 34 are used to train the image classifier. After four epochs of training with the fastai approach fit one cycle, ResNet 50 achieved a max accuracy of 98.73%. The most efficient network is ResNet 50 Mishra et al. [12]. This study recommended a methodology for determining the suitable crop and monitoring agricultural conditions in the field on a large scale using weather analysis data and crop disease diagnostics for improved production using real-time data. The data collected from sensors was input into a Decision Tree algorithm in the in-plant monitoring model to identify the plant based on its circumstances. The author used a CNN-based approach to identify and detect plant diseases from leaf images… Sladojevic et al. [13]. Berkeley Vision and Learning Center created a deep learning framework for the deep CNN training. For independent class examinations, the experimental findings on the built model achieved precision ranging from 91 to 98%, with an average of 96.3%. A new strategy involving deep learning methods was investigated in order to automatically categorize and detect plant illnesses from leaf photos. In 13 different diseases, the proposed model was able to detect the presence of leaves and distinguish between healthy and diseased leaves. The authors intend to contribute significantly to long-term development by influencing crop quality for future generations. 3000 unique images are gathered from various online sources and expanded. Sujatha et al. [14] RDL is the domain's top performer. DL works in a similar way to the human brain's neural structure, with layers and optimizers that help develop a dependable model with improved accuracy. In their work, we consider both learning methodologies, and the results of DL are noticeable when compared to ML. Precision, F1 score, accuracy, and area under the curve are all considered. In comparison, RF received a CA of 76.8%, SGD received 86.5%, SVM received 87%, VGG-19 received 87%, Inception-v3 received 87.4%, VGG-16 received 89%, and VGG-16 received 89.5%. On a regular basis, stakeholders collect and send images to the system, allowing farmers to help decide which insecticide to apply to prevent harm Mohanty et al. [15] After training a DCNN to detect 14 crop species and 25 illnesses, the trained model achieved an accuracy of 99.35% on a held out test set. This paves the way for smart phone-assisted crop disease detection and prediction. 54,306 pictures of plant leaves were analyzed. Sambasivam et al. [16] this paper proposed CNN to create a low-cost deep learning technique for detecting cassava infections, which included the following steps: Dataset collection, labeling, model training, and testing/model evaluation using k-fold cross-validation, where $k = 3$ to achieve the desired accuracy. The steps for creating the model were as follows: (a) Setting up the environment, loading, and preprocessing Data − 35% of the time, (b) Model architecture definition − 10% of the time, (c) Model training − 50% of the time, and (d) Performance estimation − 5% of the time. Ramcharan et al. [17] the author used a DCNN model to assess the applicability of transfer learning on cassava image datasets. The overall accuracy in classifying a leaf into the correct group ranged from 73% (20–70 split, knn) to 91% (20–70 split, knn) (80–10 split, SVM). The leaflet cassava dataset had higher overall accuracy, ranging from 80% (20–70 split, knn) to 93.0% (80–10 split, SVM) Ayub and colleagues [18]. To investigate the damages, the authors used data mining techniques and a family of ML techniques (Table 1).

Table 1 Summary of literature review

Ref. No#	Year	Algorithm used	Scope for improvement
1	2019	SVM, Naïve Bayes, SVM Localizer	To anticipate essential plant resistance genes and contribute to agriculture, more machine learning-based technologies are required
2	2020	RF regression model. algorithm predicted performance measurement 0.9 to 0.3	Image processing utilizing CNN to analyze current plant health can be incorporated with this system, and the results will undoubtedly be very exact forecasts
3	2020	Used regression and boosting techniques	The author concentrated on fruit diseases, but the findings can be applied to other crops
4	2019	NB method and KNN applied on rice and wheat data set	Other ML algorithms and different crops are required
5	2020	Supervised machine learning	Only experimented with one crop wheat
6	2019	SVM&RF	Only for the purpose of determining the suitability of the fertilizer to be used
7	2020	Image processing and ML Techniques	Use of DL Techniques is possible. For different types of crops
8	2020	RF	Only UsedParameter like temperature, rainfall, humidity, ph
9	2019	Biosensing mechanisms and DNA sequencing technologies	Experimented only on fungal leaf diseases
10	2020	SVM approach	Suitable to research android apps fore-commerce of product
11	2021	Preprocessing techniques performance metrics	High-quality, labeled data required
12	2021	IoT and AI	AI and IOT used in the broader sense need to be think specifically
13	2021	Computer Vision and DNN	IOT sensor may be used to collect data more accurately
14	2021	IoT and DT	Suggested to consider few more parameters to check environmental conditions

(continued)

Table 1 (continued)

Ref. No#	Year	Algorithm used	Scope for improvement
15	2016	CNN	Increasing the model's utility by training it to recognize plant diseases on larger land areas
16	2021	Proposed family of ML and DL (Inception-v3, VGG-16, VGG-19)	Used Soft computing techniques and have significant impact on the system's CA even when using a small dataset?
17	2016	Deep convolution NN for 26 diseases and 14 crop species	Improved accuracy by using more diverse set of training data set
18	2021	SMOTE with CNN	Multiple diseases co-occurring on the same plant
19	2017	Deep CNN	May be with usage of sensor and smart phone this will be more user friendly
20	2018	Decision tree, RF, NN, GNB, SVM, and KNN	hybrid approaches and DL (evolutionary and data mining)

3 Research Questions

I. Is it possible to create your own data set for predicting crop disease?
Ans Yes, Possible with the help of deploying various types of sensors and drones.

II. Is it possible to use ML/DL techniques to predict crop disease?
Ans Yes, it is possible once proper and accurate data were collected and then using machine learning algorithm, CNN, Deep learning techniques.

4 Proposed Model

Different raw data acquired through various sensors such as optical sensors, MIX sensors, motion detector sensors, and others were saved in the database in our proposed model, as shown in Fig. 1, and after adequate refinement, a data set was designed for feature selection. The disease was then predicted using proposed machine learning algorithms (s).

Phase#1: Data collections through various sensors and stores in a data base.

4.1 Optical Sensors

Optical sensors, on the contrary, use light reflection to measure and record crop and soil data in real-time crops. These crop reflectance sensors typically operate near

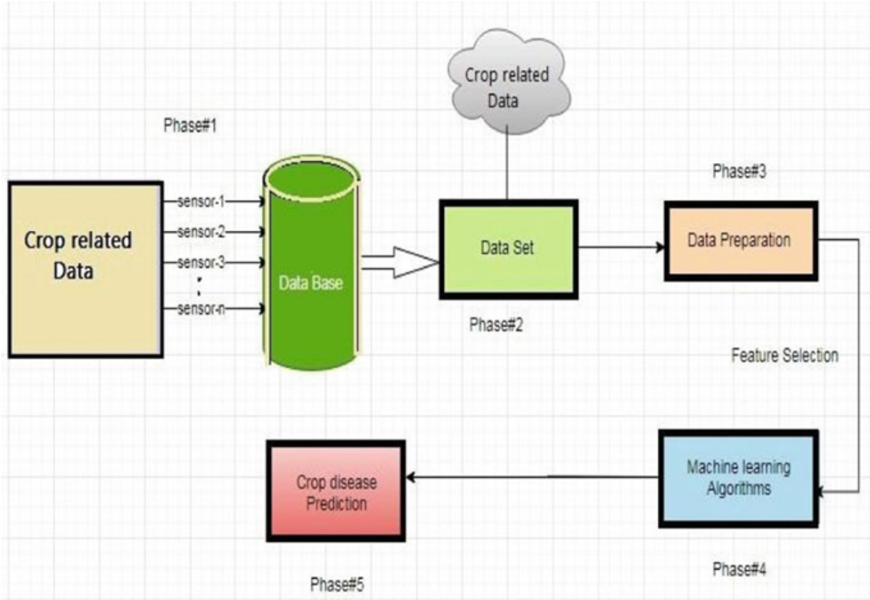


Fig. 1 FF model

infrared visible regions of the spectrum and calculate vegetation indices at least by integrating 02 wavelengths. The visible light has a direct association with the chlorophyll concentration from an agronomic standpoint, absorbing red and blue light while reflecting green light. That is why we see healthy plants as green. Electrochemical Sensors are a type of sensor that uses electricity to detect a chemical reaction. RVI is the relationship between these two wavelengths (Ratio Vegetation Index). The NDVI (Natural Difference Value Index) is a Normalized Difference Vegetation Index and the suggested normalization ensures that the NDVI values are contained in the same scale of values, ranging from −1 to 1, as indicated in the equation.

$$\mathrm{NDVI} = (\rho IR - \rho V)/(\rho IR + \rho V) \tag{1}$$

where

ρIR represent infrared reflectance;
and ρV is visible reflectance.

This sensor provides the status required for precision agriculture.

4.2 HTE MIX Sensors

It is utilized to provide moisture content and temperature in the soil as well as the surrounding environment. It is a typical environment parameter that pops up on a regular basis, and its handling is crucial in various domains. It is an electric capacity sensor that connects to a Smart Soil Moisture Sensor's cell.

4.3 Motion Detector Sensors

All across the field, motion sensors are used. When the recordings are made around the camp, those sensors can do a server-to-server transaction and then send a message to each other tool when the data is processed, all while remaining within the farm's boundaries. This device can also be used to make noise to scare animals away from potentially hazardous crops or plants.

Phase#2:

During this phase, the heterogeneous data collected from different sensors were filtered and made ready for feature selections.

Phase#3:

The process of decreasing no. of input variable(s) during the development process of a predictive model is said to be as feature selection. Its primary goal is to minimize the number of input variables in order to reduce overall modeling costs, and in some cases, it is used to improve performance.

Phase#4:

The following ML algorithms were applied to the selected features to compare and predict the disease affected to the selected crops.

Phase#5:

Most likely disease will be predicted for the selected crops, and necessary precaution may be taken to avoid major financial loss to the farmer.

5 Implementation Details

Naïve Bayes: NB classifiers are a subset of fundamental "probabilistic classifiers" based on the NB theorem and strong feature independence assumptions. They are among the most basic BN models, but they can achieve higher accuracy when combined with kernel density estimation. This theorem computes the likelihood of a subsequent event of the preceding event. Mathematically Bayes' theorem can be represented as (Table 2).

Table 2 Crop prediction using NB

Crops	Accuracy training set	Testing set
Rice	84.58	86.77
Ragi	90.32	90.12
Gram	76.59	77.77
Potato	88.22	88.19
Onion	79.77	80.25

$$P(A|B) = \mathrm{frac} P(B|A)P(A) = \mathrm{frac} P(B|A)\{P(B)\} \tag{2}$$

The above table mentioned the accuracy of the different crops. We have estimated the accuracy in terms of training as well as testing set and found NB algorithm gives 89.32 training and 90.32 and 90.12 testing accuracy (Fig. 2).

Support Vector Machine (SVM): A supervised ML technique used to solve classification and regression problems. However, this is preferably used to solve problems based on categorization types. Each data item is represented as a point in n-dimensional space (where n is the number of features chosen), with the value of each feature representing the SVM algorithm's value of a particular coordinate. After that, the classification is completed by locating the hyperplane that best distinguishes the two classes. The following is the hyperplane equation used to divide the points (for classification):

$$H : wT(x) + b = 0 \tag{3}$$

Here: $b =$ the hyper plane equation's intercept and bias term.

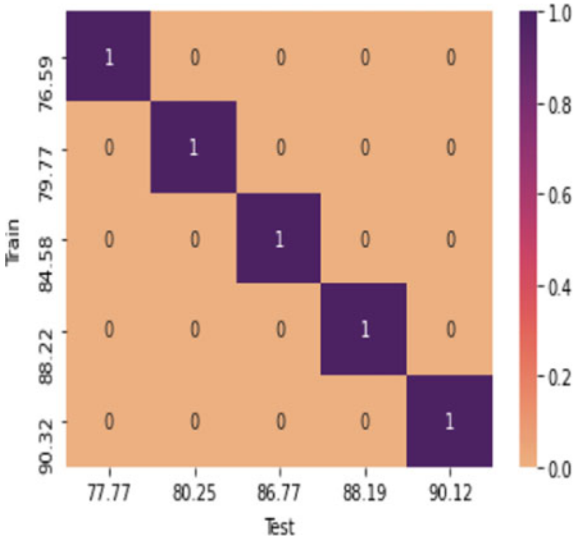


Fig. 2 Categorization using Naive Bayes

Table 3 Crop prediction using SVM

Crops	Accuracy training set	Testing set
Rice	89.78	91.47
Ragi	94.39	93.97
Gram	96.29	95.67
Potato	92.26	90.19
Onion	90.47	72.35

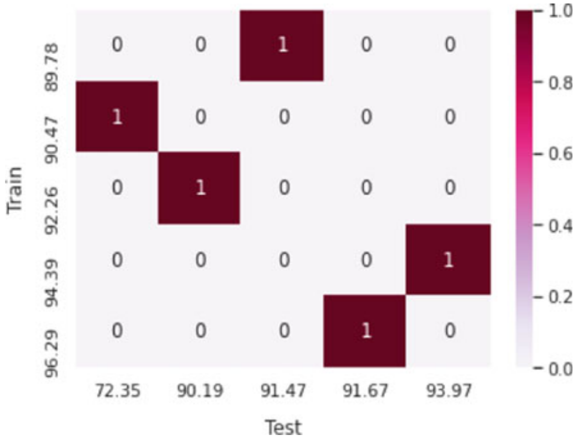


Fig. 3 Categorization using Support Vector Machine

The hyper plane in *D* dimensional space is always the D − 1 operator.

A hyper plane is a straight line in 2-D space, for example (1 − D) (Table 3 and Fig. 3).

6 Conclusion and Future Scope

Machine learning is a useful tool in today's world for analyzing massive amounts of data and producing more accurate results and predictions. Our research demonstrated and provided accurate results for the five crops that were chosen as a sample for yield prediction. Hence, we were able to confirm that machine learning algorithms may also be used to predict various illnesses impacting crops over multiple seasons and across a variety of crops because our training set and testing data are practically identical. To achieve the best level of classification accuracy, careful selection of preprocessing data methodologies and machine learning technologies is required. As a result, more machine learning-based technologies are required to predict various sorts of illnesses impacting diverse. SVM beats the other two methods, with Gram's training accuracy of 96.29% and testing accuracy of 95.67% as prediction is concerned.

More advanced technologies, such as deep learning and CNN algorithms, will be used in our future research to detect with more accuracy yield prediction and agricultural illnesses.

References

1. Yang X, Guo T (2019) Machine learning in plant disease research. Machine learning 34
2. Jawade PB, et al. (2020) Disease prediction of mango crop using machine learning and IoT. Advances in decision sciences, image processing, security and computer vision. Springer, Cham, pp 254–260
3. Akulwar P (2020) A recommended system for crop disease detection and yield prediction using machine learning approach. Recommender system with machine learning and artificial intelligence: practical tools and applications in medical. Agricultural and other industries, pp 141–163
4. Medar R, Rajpurohit VS, Shweta S (2019) Crop yield prediction using machine learning techniques. 2019 IEEE 5th international conference for convergence in technology (I2CT). IEEE
5. Sharma, Bhawana, Jay Kant Pratap Singh Yadav, and Sunita Yadav. (2020) Predict crop production in India using machine learning technique: a survey. 2020 8th international conference on reliability, Infocom technologies and optimization (Trends and future directions) (ICRITO). IEEE
6. Bondre DA, Mahagaonkar S (2019) Prediction of crop yield and fertilizer recommendation using machine learning algorithms. Int J Eng Appl Sci Technol 4(5):371–376
7. Kumar YJN, et al (2020) Supervised machine learning approach for crop yield prediction in agriculture sector. 2020 5th international conference on communication and electronics systems (ICCES). IEEE
8. Jain A et al (2019) A review of plant leaf fungal diseases and its environment speciation. Bioengineered 10(1):409–424
9. Kumari P, et al (2020) Predicition of crop yeild using SVM approch with the facility of E-MART system. Easychair
10. Fenu G, Malloci FM (2021) Forecasting plant and crop disease: an explorative study on current algorithms. Big Data Cogn Comput 5(1):2
11. Aggarwal N, Singh D (2021) Technology assisted farming: implications of IoT and AI. IOP Conf Ser: Mater Sci Eng 1022(1). IOP Publishing
12. Mishra D, Deepa D (2021) Automation and integration of growth monitoring in plants (with disease prediction) and crop prediction. Mater Today: Proc 43:3922–3927
13. Sladojevic S, et al (2016) Deep neural networks based recognition of plant diseases by leaf image classification. Comput Intell Neurosci 2016
14. Sujatha R et al (2021) Performance of deep learning vs machine learning in plant leaf disease detection. Microprocess Microsyst 80:103615
15. Mohanty SP, Hughes DP, Salathé M (2016) Using deep learning for image-based plant disease detection. Front Plant Sci 7:1419
16. Sambasivam G, Opiyo GD (2021) A predictive machine learning application in agriculture: cassava disease detection and classification with imbalanced dataset using convolution neural networks. Egypt Inf J 22(1):27–34
17. Ramcharan A, et al (2017) Deep learning for image-based cassava disease detection. Front Plant Sci 8:1852
18. Ayub U, Moqurrab SA (2018) Predicting crop diseases using data mining approaches: classification. 2018 1st International conference on power, energy and smart grid (Icpesg). IEEE

Behavioural Investigation and Analysis of Flux and Torque in Faulty Electrical Machines Using Finite Element Techniques

Hasan H. Khaleel, Amer A. Ibrahim, and Khaleel J. Hammadi

Abstract This paper presents the results of an investigation and analysis of the effects of broken squirrel-cage bars. In the investigation, a comprehensive time-stepping coupled finite element approach was fully used to compute stator current waveforms, torque, and magnetic flux density waveform. The harmonic component of air-gap flux density is analysed. From these data, the faulty signatures are extracted. The present method has been designed and implemented using Finite Elements Method depending on time stepping. The proposed method produces an efficient technique in terms of time and accuracy to detect the faults and effects on the operation of electric machine. The early detection of faults in electric machine gives enough time to decrease the probability of electric machine faults. The differences in motor torque waveform timing in each case associated with the stator current waveforms give the flux spreading in the suggested technique. The obtained results show fast fault detection and technique is founded to extract the induction motors faults.

Keywords Electrical machine · Fault's diagnosis · FEM

1 Introduction

Modern industrial of "induction motors" are frozen used drive-in manufacturing with several methods. The investigation of motor product has been considered as optimality in the usage. Such type of motor faces many problems in general manufacturing area which considers unwanted phenomena due to motor period time reduction.

H. H. Khaleel (✉)
Department of Medical Devices Techniques Engineering, Al-Esraa University College, Baghdad, Iraq
e-mail: hasan.khaleel@esraa.edu.iq

A. A. Ibrahim · K. J. Hammadi
Department of Computer Techniques Engineering, Al-Esraa University College, Baghdad, Iraq
e-mail: amer.abbas@esraa.edu.iq

K. J. Hammadi
e-mail: dr.khalil@esraa.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_2

Table 1 Specification of suggested motor

Pole number	4
Phase number	3
The outer diameter	130 mm
The inner diameter	74 mm
Air gap length	0.5 mm
Length of axial	50.6 mm
Slots number	24
Rotor bars numbers	24
Rms voltage (V)	380 V
Frequency (Hz)	50 Hz

Hence, the motor rating power monitoring decreases the probability of motor loss life. So, the requirements of "faults diagnosis" in this type of motor could be achieved by many advanced techniques [1]. Then, the harmonic current and other many parameters in the motor stator will affect the motor efficiency due to its increasing and decreasing compared with standard values [2]. Any of these parameters could be used as useful signs of faulty motor "diagnosis" of induction motors. The finding of these catalogues in many ways could suggest solving this problem [3]. The potential magnetic aberration parameters and the spreading of flux are used to overcome many faults in the broken bars of motors [4]. A problem is protecting the motor device indoors from the "air gap" which affect the incoming signals from motor. To sense the rotor bar faults, the turning radius techniques could be used [5, 6]. Consequently, this technique needs more additional devices. Many researchers introduce advanced methods to diagnose the motor faults as in [7–9]. One of the more attractive methods in these fields is the consideration of harmonic produced by stator current as in [10, 11]. The investigated method of [12] which used the "inverted fed induction motors" caused an unwanted frequency harmonic. The idea by [13, 14] of in hand fault motor using frequency spectrum illustrates peak power of spectrum and pole number in the motor consideration. The opera-2d programme technique is carried out to sign the faults by means of broken bar rotor routine. The specifications of suggested motor are illustrated in Table 1.

2 Mechanical Fault Concept

The common things affected on rotor could be simplified and explained briefly in the healthy bars, which are forced to carry high current and high-speed torques. Additionally, the large air in "aluminium" causes uniform bars to expand. The static and dynamic rotor symmetry could affect rubbing the stator winding.

3 Mathematical Model

The most techniques used to model the fault diagnosis and investigate the induction motor damages are the "winding function, equivalent circuit and finite element methods". The high performance of electromagnet symmetrical in the induction motor could be used in the suggested model, which indicates any asymmetry degree sign as illustrated in Fig. 1.

An asymmetrical of electromagnetic field in the motor produces inequalities in the phase resistance as well as in the air-gap electromagnetic field which introduces a harmonic frequency in the current of stator and rotor. This phenomenon could be expressed in the mathematical model which adds extra resistance to phase of rotor as in [15].

$$R_{\mathrm{bar}} = \frac{l_{\mathrm{bar}}}{\cos\alpha_{\mathrm{skew}}.\sigma s} \tag{1}$$

$$\Delta r_{ra,b,c} = \frac{3n_{bb}}{N_b - 3n_{bb}} r_r \tag{2}$$

where $\Delta r_{ra,b,c}$ phase change of rotor resistance.

The change produced by rotor effects $\left(\Delta r_{ra,b,c}\right)$ is created by rotor effects which are derived depending on the statement of broken bar as touching either resistance of ring end or magnetizing current. A corresponding resistance of healthy machine could write as:

$$r_r = \frac{(2N_S)^2}{N_b/3}\left[r_b + \frac{2}{N_b\left(2\sin\frac{\alpha}{2}\right)^2} r_e\right] \tag{3}$$

where

r_b and r_e denote the end-ring resistances and rotor bar,

N_s turns of winding.

To simplify Eq. (2), assume r_e deserted, then Eq. (2) become:

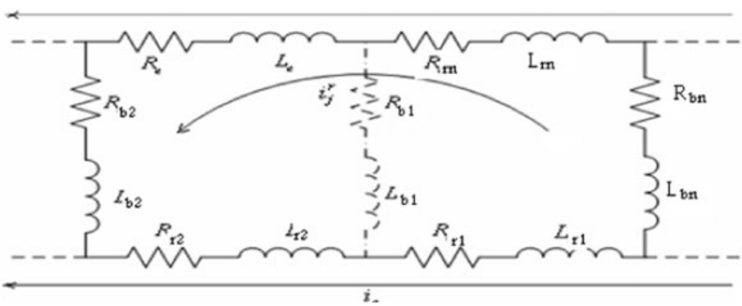


Fig. 1 Allocation structure

$$R_{\text{ring}} = \frac{\mathrm{l}_{\text{ring}}}{\cos\alpha_{\text{skew}}.\sigma s} \tag{4}$$

$$r_r \approx \frac{(2N_s)^2}{N_b/3} r_b \tag{5}$$

and

$$r \approx \frac{(2N_s)^2}{\frac{N_b}{3} - n_{bb}} r_b \tag{6}$$

Hence

$$\Delta_r = r - r_r = \frac{3n_{bb}}{N_b - 3n_{bb}} \tag{7}$$

The 2nd measurable fault assessment as suggested by [16–21] is:

$$\frac{I_{bb}}{I} = \frac{\sin\alpha}{P(2\pi - \alpha)} \tag{8}$$

which I_{bb} and I *are* amplitudes of the sideband frequency α …… electrical angle.
Therefore, the electrical angle could be expressed as:

$$\alpha = \frac{\pi P n_{bb}}{N_b} \tag{9}$$

4 Tests and Results

The finite element method approach is used to diagnose the effect of broken bars in three-phase induction motor which achieves computing healthy and faulty cases and collects the data as waveforms for magnetic field and air-gap flux. This data could extract different fault type signs without destroying the motor. The programme of opera-2d is used in this work to observe and detect the faults and provide good proof of suggested approach. The plots of rated motor in the case of asymmetrical and symmetrical rotor condition are presented. Hence, a flux density of healthy motors shows a bit unbalance due to rotating field introduced by faulty motor as shown in Fig. 2.

Figure 3 illustrated a current through stator in faulty and healthy state under load condition. The stator current increased suggestively associated with a fine rotor because the component of harmonic which caused by faults is superimposed into stator current winding at the frequency f_b. Concentrating faults on one pole produce

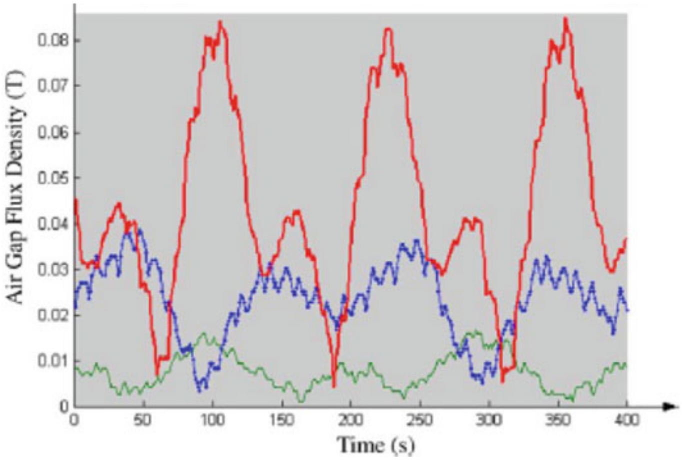


Fig. 2 Waveform of flux

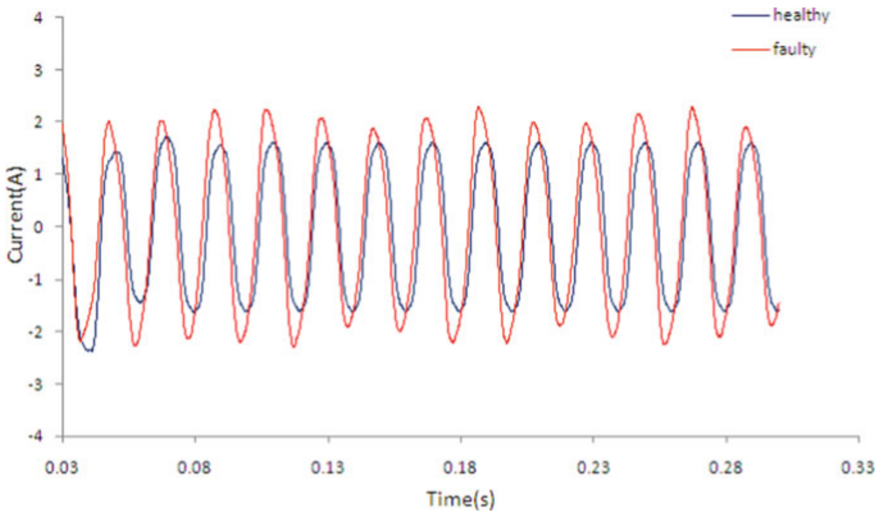


Fig. 3 Variation of current

an increase in the value of harmonic mechanisms because the faults cause irregularity of flux spreading, that produce more harmonics. Faults generate amount of frequencies with the rotating flux waveform. The rotating flux waves could encourage currents in equal frequency with the stator current.

The difference between two stator amplitudes is due to fault case because of asymmetrical flux distribution case, more harmonic in the faulty motor performance. Hence, increasing the number of broken bars would introduce bad asymmetrical in flux. High rise in the frequency harmonic will appear in the case of increasing the

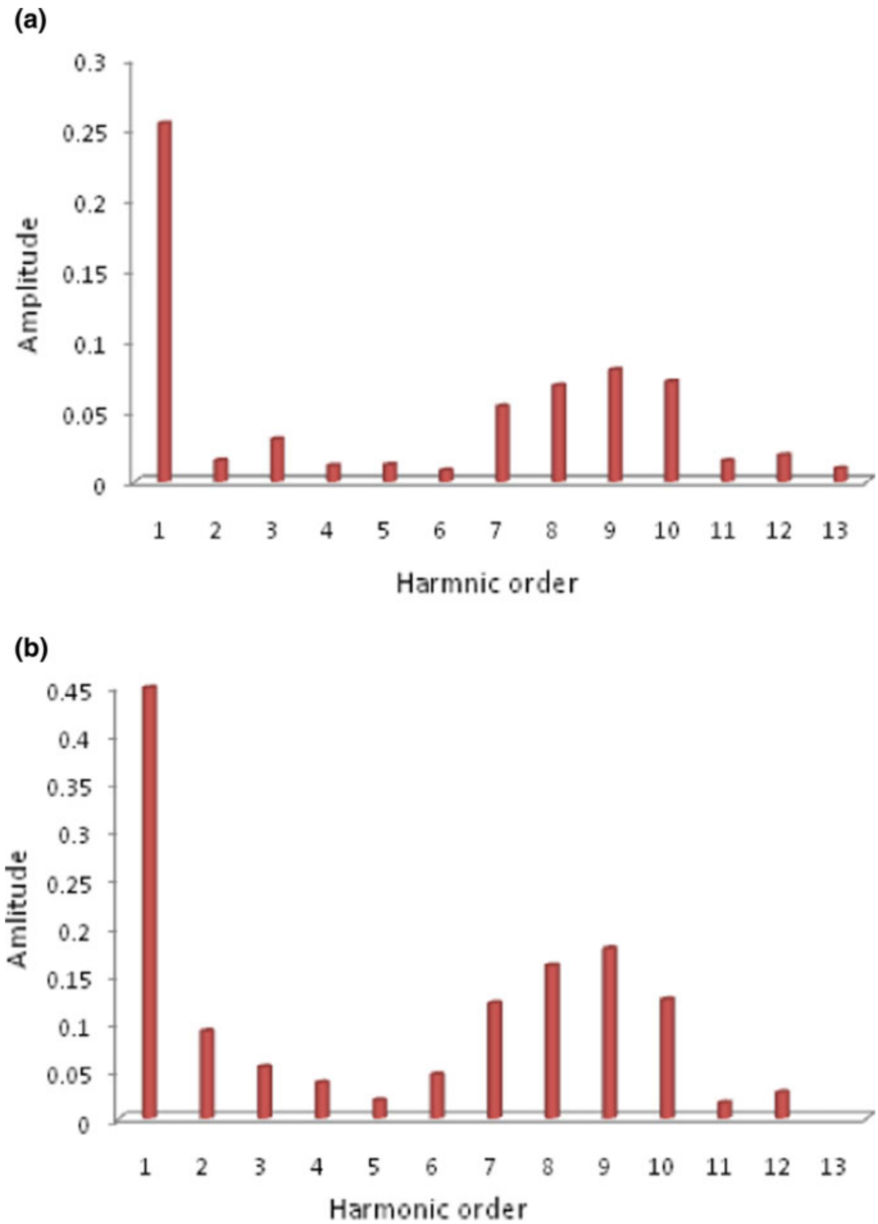


Fig. 4 Harmonic spectrums

amplitude of harmonic of flux waveforms. As a result, one could observe that the harmonic in the faulty motor appears more asymmetrical compared with healthy

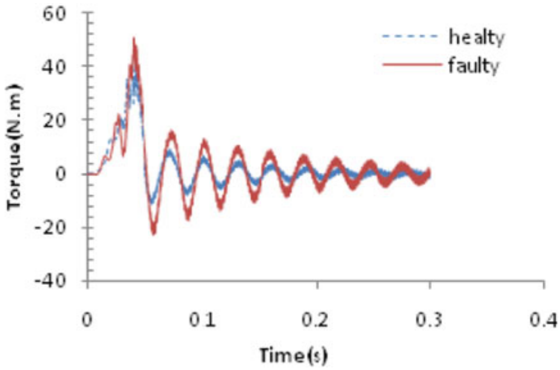


Fig. 5 Variation of torque

motor. Figure 4 illustrates the air-gap harmonic spectrum in healthy and faulty machine and the difference between them is highlighted as well.

In the faulty machine, the waves of magnetic torque are more compared with healthy motor. The time variation of the torque in the healthy and faulty more is illustrated in Fig. 5. The oscillation of developed torque is increased in case of faulty motor more than traditionally enveloped in the healthy motor as well more values. The time variation of torque shape in faulty motor is seen more than in healthy one and the ripple is high also.

5 Conclusion

This paper presents the detection of mechanical faults in induction motor by using the opera-2d programme techniques. This approach will increase the time life and keep the motor efficiency as top as possible. The early detection of the faults in the motor will give enough time to protect the machine and decrease the cost by many phases of running time. The suggested technique provides and enhances the machine researcher to develop the fault detection in the current and future generations of induction motor. This work opens a new window for developing and accelerating the induction motor design process to get efficient motor. The future direction of this study aims to develop a comprehensive database of non-invasive sensor measurements of induction machines with different faults and to understand the effect of faults on each sensor output.

References

1. Nandi S, Bharadwaj R, Tolivat HA, Parlos AG. Study of three phase induction motors with in cipient rotor cage faults under different supply conditions. In: Proceedings of the IEEE-IAS 1999 annual meeting, vol 3
2. Faiz J, Ebrahimi BM, Sharifian MBB (2006) Different faults and their diagnosis techniques in three phase squirrel-cage induction motors—a review. J Electromag 26
3. Paoletti GL, Rose A (1989) Improving existing motor protection for medium voltage motors. IEEE Trans Ind Appl 25(3)
4. Fiser R (2001) Application of a finite element method to predict damage induction motor performance. IEEE Trans Magn 37(5)
5. Povinelli RJ, Bangura JF, Demerdash NAO, Brown RH (2002) Diagnostics of bar and end-ring connector breakage faults in poly phase induction motors through a novel dual track of time-series data mining and time-stepping coupled FE—state space modelling. IEEE Trans Energy Convers 17(1)
6. Bangura JF, Povinelli RJ, Demerdash NAO, Brown RH (2003) Diagnostics of eccentricities and bar/end-ring connector breakages in poly phase induction motors through a combination of time series data mining and time-stepping coupled FE—state space techniques industry applications. IEEE Trans Ind Appl 39(4)
7. Faiz J, Ebrahimi BM, Sharifian MBB (2007) Time stepping finite element analysis of rotor broken bars fault in a three-phase squirrel-cage induction motor. Progr Electromag Res PIER 68
8. Mirafzal B, Demerdash NAO (2005) Effects of load magnitude on diagnosis broken bar faults in induction motors using the pendulous oscillation of the rotor magnetic field orientation. IEEE Trans Ind Appl 41(3)
9. Ondel O, Boutleux E, Clerc G (2006) A method to detect broken bars in induction machine pattern recognition techniques. IEEE Trans Ind Appl 42(4)
10. Haji M, Toliyat HA (2001) Pattern recognition—a technique for induction machines rotor broken bar detection. IEEE Trans Energy Convers 16(4)
11. Bangura JF, Demerdash NA (1999) Diagnosis and characterization of effects of broken bars and connectors in squirrel-cage induction motors by a time-stepping coupled finite element—state space modeling approach. IEEE Trans Energy Convers 14(4)
12. Innes AG, Langman RA (1994) The detection of broken rotor bars in variable speed induction motor drives. ICEM94
13. Legowski SF, Trzynadlowski AM (1995) Instantaneous stator power as a medium for the signature analysis of induction motors. IEEE IAS 95
14. Trzynadlowski AM, Ghasezadeh M, Legowski SF (1997) Diagnostics of mechanical abnormalities in induction motors using instantaneous electric power. IEEE, MBI 9.1–9.3
15. Bentounsi A, Nicolas A (1998) On line diagnosis of defaults on squirrel cage motor using FEM. IEEE Trans Magn 34(3):3511–3514
16. Theodore wildi (2000) Electrical machines, drives, and power systems, fourth systems, Tk2182. W53
17. Stephen J. Chapman British aerospace Australia, electric machinery fundamentals, 3rd edn. MCGRAW-HILL
18. Tata McGraw-Hill. Electric machines, 3rd edn. Publishing company limited (New Delhi), 3rd reprint 2005 RYLQCRABRCAQZ
19. Pyrhonen J, Jokinen T, Hrabovcova V (2009) Design of rotating electrical machines, Ltd. Wiley, TK2331, p 97
20. Workshop in Design and modeling of electrical machine, school of electrical and electronic engineering USM engineering campus March–2009 By Engr. DR. Dahaman Isahak
21. Regin R, Obaid AJ, Alenezi A, Arslan F, Gupta AK, Kadhim KH (2021) Node replacement based energy optimization using enhanced salp swarm algorithm (Es2a) in wireless sensor networks. J Eng Sci Technol 16(3):2487–2501

22. Ho SC, Hong CG, Hwang GJ (1989) Transient and steady state performance of a squirrel-cage induction motor. IEEE Proc Part B 136(3)
23. Fišer R, Makuc D, Ambrožič V. Evaluation of the induction motor cage fault stage using finite element method. Record of IEEE SDEMPED'2001. The 2001 IEEE international symposium on diagnostics for electrical machines, power electronics and drives, Grado (Italy), pp 627–631
24. Charles I, Hubert PE (2002) Prentice hall, electric machines, theory, operation, applications, adjustment, and control, 2nd edn. TK2182. H83

A Comprehensive Survey for Internet of Things (IoT)-Based Smart City Architecture

Rohit Sharma and Rajeev Arya

Abstract With the advent of mobile technology, the modern paradigm of "connected everyday objects" was built over the current network. The tremendous development of networked devices had increased its reach over the primitive network topologies. This significant change has launched the revolution after flat-page. The surge in the global urban population is placing new demands on people's daily lives in terms of pollution, public safety, road congestion, etc. To accommodate this rapid growth, new technologies are being developed and smarter cities are being built. Incorporating the Internet of Things (IoT) into everyday life makes it possible to develop new smart solutions such as services and applications for industries like hospitals, surveillance, forestry, etc. Research on Artificial Intelligence (AI), Deep Learning (DL), and help of Data Visualization have shown how IoT performance can be improved with some technological aids. This creates a rapid demand for addition and works in terms of Big Data with first-class technologies that we have around us, so in this paper, we will talk about such things and show a comparison on this basis with the other works that are under it, with the deep learning and artificial intelligence models. This study will help us to show how it overall contributes to the growth of the Internet of Things in society to provide a better life for future generations. Finally, we will outline the existing obstacles and problems that occur during the smart city growth facilities.

Keywords Internet of things (IoT) · Deep learning (DL) · Artificial intelligence (AI) · Technology · Services · Applications · Smart city · Innovation · Data visualization · Big data

R. Sharma (✉) · R. Arya
Wireless Sensor Networks Lab, Department of Electronics and Communication Engineering, National Institute of Technology Patna, Bihar 800005, India
e-mail: rohits.ph21.ec@nitp.ac.in

R. Arya
e-mail: rajeev.arya@nitp.ac.in

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_3

1 Introduction

The Internet of Things is a spider web of software, sensors, databases, and other devices superfluous web. The hardware differs in intricacy from ordinary household chattels to suave business systems. Analysts say that by 2020 there will be a shift to 10 billion connected IoT gadgets, and by 2025 there will be an upward growth of 22 billion [1]. IoT has emerged way lot over the past few decades and gaining importance day by day. This helps us in day-to-day living as we have multiple gadgets interfaced using the Internet which makes our life easier in simple terms. Digital systems are very efficient in log management and also sorting life in a better way with the aid of computing technology and analytics in providing a solution to mankind's problems. The IoT is a distributed web of connected gadgets, which are Internet-enabled entities that can receive and send information non-connection oriented without human interaction. These possibilities for reserved or business growth exceed boundaries. A networked biological device, a small electronic chip transceiver, a solar cell, a networked automated braking equipped car (gas, tire-tubes pressure, required scheduled maintenance, and much more), or other item equipped with electronics and the ability to collect and emit data wirelessly is a "thing". Industrialists are now taking inspiration from IoT and developing positive revenue, lower operating costs, and better customer service, as well as increasing efficiency in regulatory enforcement, are also driving businesses. IoT system implementations provide the bulk datasets and inner results needed to optimize workflows, visualizing user patterns, automating the processes, meeting parameters, and operating more effectively in a changing marketplace for whatever reason. The IoT is an ever-growing network capable of autonomously discovering and sharing data about distinct devices that are special. Due to the success of embedded tech and the exponential growth of embedded device technology, the IoT has become the focus of various stakeholders.

Recent developments in a variety of technologies have made it a reality. Less-cost, less-power sensor tech is there. IoT technology is now getting the door-to-door and suppliers and industries due to its less cost and varied applicability. Cloud computing platforms and cloud services are becoming more widely available, giving businesses and customers access to the technology they need to scale without having to worry about everything themselves. IoT has paved the way for big data analytics and the data collected through these sensors is turning into a mine of gold for companies and business is taking place as a result and they are selling these data for better and efficient life by analysing all the odds. Artificial intelligence (AI), conversational, and natural language processing (NLP) have made their way into IoT devices (e.g. Google, Bixby, Alexa, Cortana, IBM Watson, and Siri digital personal assistants) thanks to advances in neural networks, making them more attractive, accessible, and viable for home use. So, from the above interpretation, we can conclude that IoT means connectivity between billions of IoT-based devices. Such capabilities give power to normal gadgets that can run in real-time, adapting with conditions, and run with zero interference/control. Electronics which are now embedded in various urban centres, resulting in a huge amount of data being collected. Before it can be viewed, the collected data needs to be stored, analysed, and processed using specialized models and visualization tools to make it useful for better insights. The recent emergence of IoT data has led IoT big data analytics, including machine learning analytics, to gain significant traction recently: ML, DL, and Computing Infrastructures, especially as conventional data processing techniques have encountered numerous limitations, especially when dealing with big data. DL algorithms have paved way to deal with similar problems. DL algorithms enable developers and enthusiasts to process real-time data with great precision and relatively much better performance. Plethora of studies have been conducted with DL. The results were ranging from various domains which led a broader sight for IoT-based cities and other baby strides towards their growth. In this report, we have chosen Smart City as the primary application because IoT-based cities have wide instances and case studies, the IoT sector instances. We then describe instances that arise in a speculative IoT-based city, including in home, in health care, in transportation, in surveillance, in agriculture, and in climate monitoring system. We further did a juxtaposition based on several parameters with challenges faced in these cities. Apart from this, we tried covering out the recent trends and issues going round the globe.

2 Related Work

Atitallah et al. [2] have presented a survey paper over the IoT-based cities, and they include a summary of the literature on the use of IoT and DL to build smart cities in their survey. They start by defining the Internet of Things (IoT) and describing the characteristics of IoT-generated big data. Then, they go through various computing infrastructures that are used for IoT big data analytics, such as cloud, fog, and edge computing. Following that, they review recent research that uses both IoT and DL to create smart applications and services for smart cities, and they survey common DL models. Finally, they have discussed some of the existing problems and concerns that have arisen as a result of the growth of smart city services. Silva et al. [3] worked on big data analytics for embedded smart city architecture for estimating the performance and worked for its enhancement through real-time data processing and decision-making for data collected for a year through their hardware and resources. The integration of Big Data analytics with smart city architecture is discussed in this paper to suggest a practical and feasible structure for the implementation of smart cities. Real-time intelligent decision-making, autonomous data collection, and user-centric energy customization are all possible with the proposed architecture. The most influential factor for the realization of a smart city, however, is decision and control management. As a result, the new scheme's most important aim is to achieve real-time and timely decisions. Fusion techniques also help to speed up the analysis of the massive amounts of data generated in Big Data analytics. Mohammadi et al. [4] did a comprehensive study on the different areas including in home, in the city, in energy, in healthcare, in Agriculture, in Education, in Sports, in Retail, and in IoT-based infrastructures, for IoT applications and facilities, with the most up-to-date deep learning (DL) methods which are tested. Side by side with the analysis using Big Data Analytics Model, Mahdavine et al. [5] in this survey investigated various types of ML techniques and methodologies implemented to probe and summon IoT-based data in various IoT-based smart city use cases. The writer's defined and gave a juxtaposition of ML algorithms and reviewed the usage. Also, they discussed the vulnerabilities and challenges at last which are implemented in big data analytics. The significance of this paper's contribution. The aim is to respond to the following questions: What kind of MLA could be used on IoT smart data? What is the classification of MLA that can be used in the Internet of Things? What is the facet of IoT data in the real world? What makes the IoT-based City a standard IoT technology use case?

Zhang et al. [6] in their paper discussed the different wireless networks typically in mobiles with the implementation of DL. Also, some basics of deep learning with it for better understanding for the readers highlighting immediate problems and advantages in mobile network and a study based on the different aspects and scenarios. Zhang et al. [7] in their research studied four different ways by which DL is used in big data which are as follows: stacked auto-encoding method, deep belief networking method, CNN, and RNN. Also, they made some classifications based upon their work: bulk data group, distinct data group, the streaming data group, and finally the reduced-quality data group. Qolomany et al. [8] in their paper studied and showed how different ML is used in big data to make a qualitative analysis. They reported on and classifications made in IoT-based smart cities on various parameters which helps a city to run and depend on each other at different topologies. Chen et al. [8] conducted a brief study on how DL is used in smart IoT-based cities. He looked at the most common models and summarized the most recent researches in various IoT-based cities using cases and instances, such as vehicle management, the healthline sector, the climate, and data privacy.

3 Recent Trends and Overview

In the Figs. 1 and 2, we can see a graph which depicts the search interests of people around the globe for Internet of Things from year 2004—present; the graph has increased tremendously. This field is gaining more and more importance in today's world due to its applicability's in daily life and easing everyone's task.

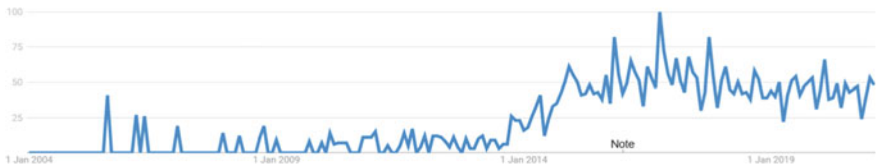


Fig. 1 Internet of Things trends search history interest on google search results

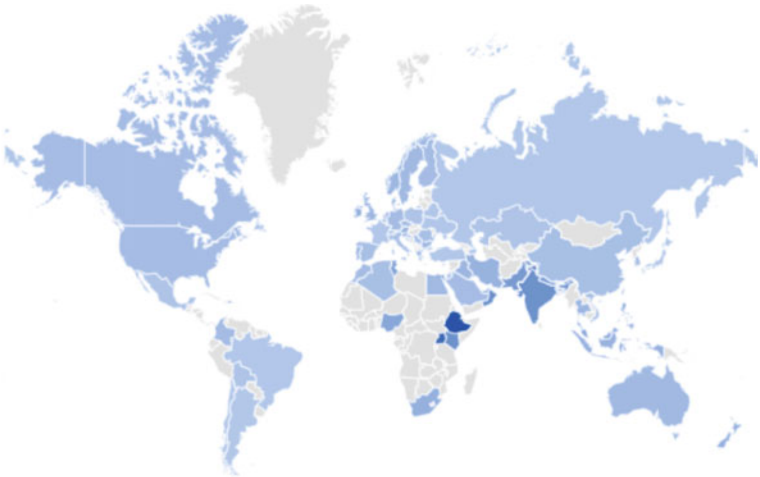


Fig. 2 Worldwide trend for Internet of Things as per search history interest on google search results

Thus, more and more people are engaging for research and finding more application of IoT. The second graph it is shown on the world map over 7 continents showing peoples interest over these many years. The darker the region, the more the search interest and people's engagement with that topic and the greyed-out areas still have not searched it on the Internet [9]. The time interval between each year is taken around 5. Out of which, if we talk about the ranking of top five countries based on their search interests then Ethiopia ranks first, Mauritius ranks second, Singapore ranks third, India ranks fourth, and Uganda ranks fifth, with having related search queries topics as the Internet of Things, cloud computing, artificial intelligence, big data and what is Internet of Things? etc [10].

In the above Figs. 3 and 4, we can see a graph which depicts the search interests of people around the globe for big data analytics from year 2004—present the graph in comparison to Fig. 1 has increased violently and much sharper and have been increasing [11]. Just because after the infrastructure have been implemented for IoT the work does not stop here the hardware generates a large amount of sensitive data year after year which needs to be stored somewhere for which we use cloud databases and people and this new field of IT is emerging at a lot faster pace and the demand for data scientists and engineers is increasing who can study this big amount of data and

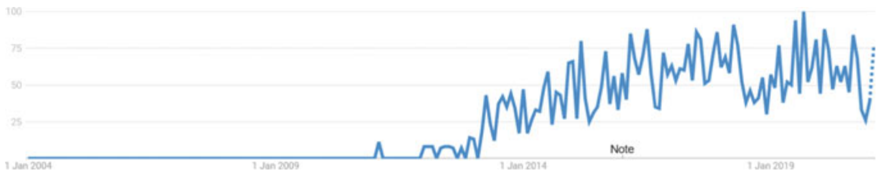


Fig. 3 Big data analytics trends search history interest on google search results

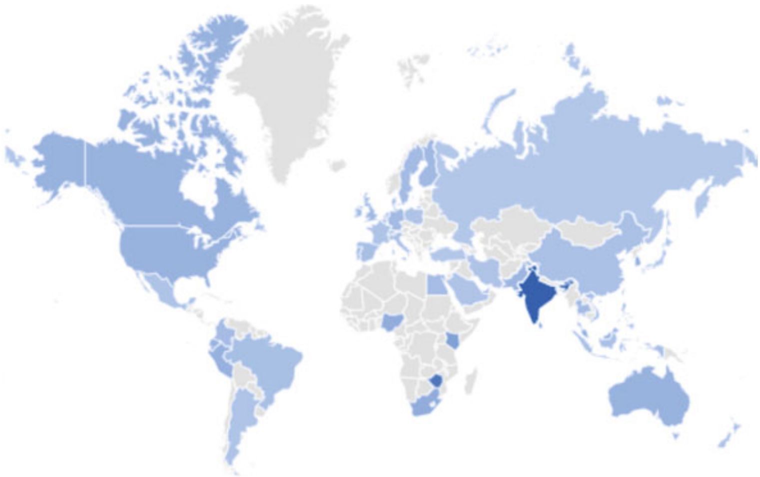


Fig. 4 Worldwide trend for big data analytics as per search history interest on google search results

can provide efficient results for the mankind to improve the quality of our lifestyles. The second graph: It is shown on the world map over 7 continents showing peoples interest over these many years. The darker the region, the more the search interest and people's engagement with that topic and the greyed-out areas still have not searched it on the Internet also can be termed as no work is possibly been done here on this topic and field. The time interval between each year is taken around 5 [12].

Out of which if we talk about the ranking of top five countries based on their search interests, then Singapore ranks first, India ranks second, Zimbabwe ranks the third, Hong Kong ranks the fourth, and Sri Lanka ranks the fifth, with having related search queries topics as big data analytics tools, big data analytics salary, tools for big data analytics, best big analytics tools, data analysis, etc [13]. Also, we can notice a difference in between Figs. 1 and 2 that Internet of Things took popularity around the year 2007, and whereas the term big data analytics gained importance at quite a later stage, i.e. around 2011 when it was felt that there was a need to secure this large data and, also, we can study this data for providing efficient solutions to the society. On the contrary, the first IoT-based device was discovered in early 1980s and the big data term was coined around 2005 [14].

In the Fig. 5, it shows an analysis of total estimated technological investments worldwide that is to be made in billion dollars which is estimated to be around 189.5 billion US dollars approx. towards it by the year 2023. By this graph, we can also conclude that the amount of money invested in this field has increased gradually since 2018 to 2023 [15]. In our paper, we will focus on some sectors of smart city to check onto the different technologies that are used in these and who these are integrated all together. Which is shown in the below flowchart Fig. 6.

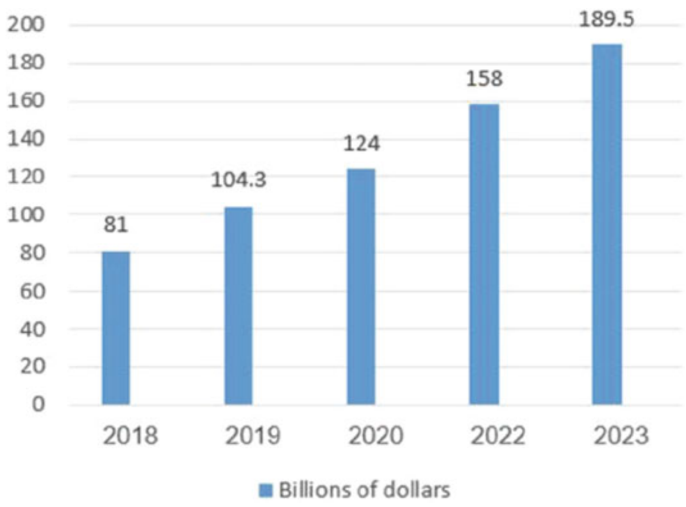


Fig. 5 Prediction made for technological investment for IoT worldwide

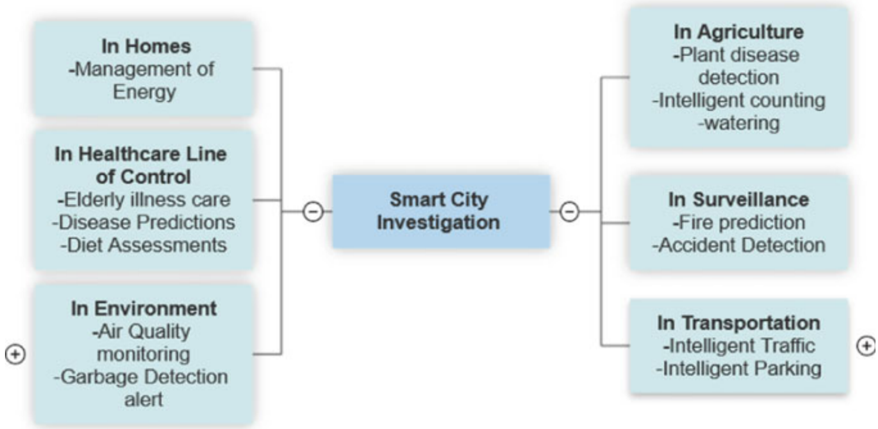


Fig. 6 Smart city investigation chart

In homes, we know that IoT is implemented for energy management; apart from this, it is also used for audio systems, security, inventory management, etc. In agriculture, IoT is used for detection of plant disease, and water regulation for watering the plants and other applications like soil Ph level and pesticide tracking. In healthcare department, it is implemented for the treatment of people above the age of 60, prediction of some hard-to-find disease using deep learning and artificial intelligence also, nowadays robotic surgery is happening which is responsible for zero human interaction, and someone can operate from far off place. In surveillance, it is used for predicting the outbreak of fires in buildings, industry or public places, as well as accident management [16]. In environment, IoT is responsible for monitoring the air quality and it can be done both indoor or outdoor to determine the colloidal or suspended particles in air which can be an issue for asthma or breathing-related disorder patients, as well some gases in environment and weather prediction as well based on the humidity level in air and wind movement and garbage detection as well. In transportation, IoT is very much helpful easing people's life by traffic management, parking management, and also estimating time of travel and various other things which help us.

4 Technologies Involved

4.1 Architecture of IoT

Before we discuss about the technologies involved in smart cities based on IoT, let us talk first about its architecture. IoT's architecture see Fig. 7, involves hardware which usually comprises sensors, printed circuit boards, Internet connection, communication channel, Hadoop system for storing data as a data lake, and analytics based on the data collected from different data marts based upon the criterion decided by the Hadoop developer and Data analyst [17].

Here, for storing data people prefer NoSQL databases or cloud databases to store data which are large and raw data which cannot be used directly. The analysis

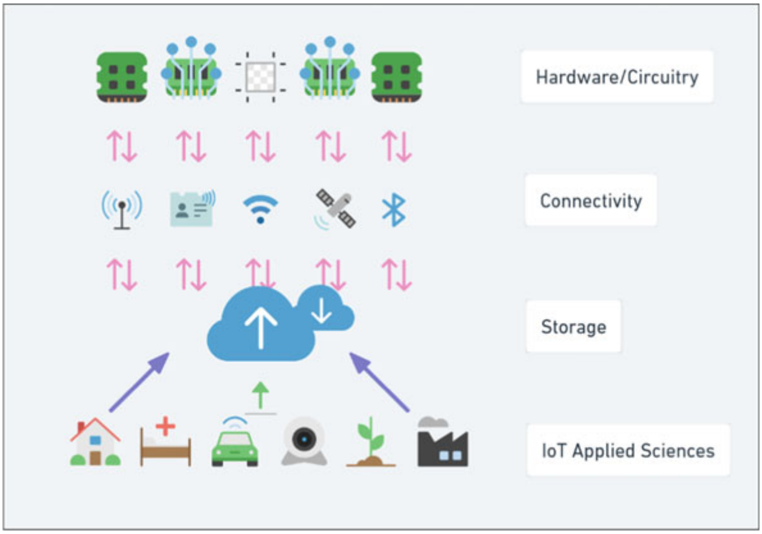


Fig. 7 IoT Architecture

performed over these data after filtering out is of three kinds: descriptive analytics, predictive analytics and prescriptive analytics (which is discussed later this section). This helps us to get an estimate or exact data for developing applications and software for them. As data analysis is performed over a history of data, thus the more accurate results. The examples shown in Sect. 3 for search trends have a data from 2004 to 2021, thus showing a variation over a long period of time [18].

4.2 Analysis

In the **descriptive analysis**, there is a historical data which actually shows a juxtaposition between different data over a period of time such type of analysis is used mainly in industries for comparing their growth, finding potential customers, etc. In the **predictive analysis**, this is used by processing data with some BI (Business Intelligence) tools and performing ETL operation (Extract, Transform, Load) or maybe an ELT (Extract, Load, Transform) thus, as a result finding us a pattern over a range or sequence of data. In the **prescriptive analytics**, this method is a successor and a much more intelligent method over the said above this method actually analyses the data and before giving any kind of decisions it gives us the exact results which helps us to take decisions much better way [19].

4.3 Cloud Computation Techniques and Concepts

Now, we will see computational techniques the different **Cloud Computing**, technologies in big data analytics; the basic definition for the same is the data which can be accessed from the Internet form anywhere at any point of time and any number of times. Some cloud computing services which are offered nowadays are based ***storing data over the Internet***; in these kinds of servers, we use TCP/IP connections for the data pipelining. Others, like ***services that are featured over the Internet*** , in these kinds of cloud computation technique all the ready-made options are already there and we are supposed to perform our operations on them and get the result. The last one is the ***apps that run on the Internet***, and these are the files or software that run on Internet which are installed and often come up with some GUI.

Whatever data is collected from the systems installed, goes to processing which is based on deep learning algorithms designated specifically to find the results. Like, in our project we used one such model for end-to-end deep learning over the data for analysing. But these include some limitations such as like **cost of infrastructure** is too high for running such model and frameworks, signal strength is too important else **latency** can be a disaster while data analysis, redundant infrastructure should be there and some disaster management is required and data replications should be stored in racks and no more than two replicas should be stored in a single rack. Also, **data privacy** is one of the essential factors in this case as we have huge amount of sensitive data which can be used by people who have malicious intent [19]. Other Computing techniques are Fog and Edge Computing. In *Fog Computing,* Fog is a component of something like a distributed network environment which is correlated to cloud computing and the Internet of Things (IoT). Computing resources that provide public infrastructure as a service could be thought of as a large, global destination for data; data from IoT nodes is formed at the network's edge. Fog connectivity is a portion of the Internet of Things (IoT) idea, which anticipates most of the device's individual use on a daily basis being linked to one another. Smart phones, wireless health tracking technology, connected vehicles, and VR technology using devices like Smart Glasses are only a few instances. *Edge computing* is an emerging virtualization technology that pushes business processes closer to datasets like IoT systems and local edge servers [20]. This relative vicinity to data at its source does provide large financial advantages, such as better perspectives, faster response times, and also more spectrum accessibility.

4.4 Models and Tools Used

Artificial Intelligence as term itself says that it is a kind of intelligence which is artificial not god's creation it is manmade creation which is able to compute and process data and provide us solution to daily life issues or business problems. This comes under Deep Learning; due to this feature, we are able to teach machines and learn new tricks from their errors using machine learning algorithms. There are many Deep learning Algorithms which are used in smart cities; for data analysis, we will discuss Deep learning models below:

- Yolo v3
- CNN and Mask R-CNN (Convolutional neural network.)
- RNN (Recurrent neural network)
- DBN (Deep belief Network)
- SAE (Stacked auto encoder)
- GAN (Generative Adversial network)
- DRL (Deep reinforcement learning).

These models can be further categorized into four types of techniques: supervised deep learning model, unsupervised deep learning, semi supervised deep learning and reinforcement deep learning.

Some tools which are used in deep learning (DL) are as follows:

- TensorFlow—it is a kind of framework for dl models interfacing and can be done by python, C++, Java, and Go language.
- Torch (if in python use pytorch lib)—it is also a framework for running dl models; it is interfaced via C, C++, Lua, OpenCL
- Theano—it is python-based library used in low level apis, and this is used mainly for RNN

- Caffe—it is a framework used using C++ via python, C++, MATLAB.
- Keras—it is library in python used in high level apis
- MatConvNet—it is MATLAB-based toolbox which is written in Cpp language
- MXNet—it is a library written in Cpp and often used using python, R, Scala, Cpp, Perl
- Google Colab—it is written in python and its best suited for machine learning, data analysis work.

5 Methodologies of Dl in IoT-Based Smart Cities

Now we will focus on how these deep learning methodologies are used in smart cities.

In Homes: Nowadays, most of the urban and modern homes are connected with Internet at every corner of house, thus with the advancement, people want everything at their palms to control things around which is possible through IoT they control all these using smartphones. Appliances like lights and bulbs, air conditioners, cameras, fridge, kitchens, or personal assistant robots and many other such things here, we can see how these things are connected together to give the optimal experience to the user and live life in a luxurious way. In smart IoT-based homes, the use of LSTM and CNN can help us in regulating the energy and managing the load-based AI. The data collected from various sensors are collected and stored in a db which acts as a data lake from here the deep learning algorithms designed takes the data and show the result and predictions; such results are used by companies for designing better and improved sensors and applications for future generations [21].

In Health care: In this sector, IoT is like a boon to mankind and helps in saving lives of many, and these are available in the form of some gadgets which can be wearable or not wearable, some trackers, or some transducer in bio medical which checks for some ailment and reports it into a db or AI-based system. This many a times helps us in estimating the future diseases and many other things which can be inhibited at a very early stage. One such example recently is neuralink which is developed by a neuro tech-based company which is founded by Elon Musk, this chip is implanted in human brain and brain can be connected through a connectionless medium to a smart phone which helps us to track many activities till now it has been experimented in pigs and have given successful results. Also in health care, OPEN CV, YOLOv3, LSTM, CNN, and GRU models are used for different applications such as treatment of people above the age of 60, prediction of chronic and acute diseases in people and kids. It is helpful in predicting some genetical disorder which is nowadays, a part of research in biomedical engineering. Also, this is helpful in creating reports collected from blood samples, or predicting intolerance charts in diet planning. Also, in estimating if a person is COVID positive based upon the symptoms ensuring social distancing and create a huge database and record and predictive analysis to control the wide spread of disease and create a vaccine [22].

In Environment: These sensors implemented at different institutions and buildings can also be used for applications where there is need to monitor environment such as air quality index and humidity. Weather conditions and wind speed, also, can be used for garbage management. RNN, ANN, CNN, R-CNN, and LSTM are used for estimating and prediction; this is also beneficial for prediction of pressure prone areas and forest fires as well. Thus, we can find a solution beforehand before any major loss.

In Agriculture: The use of IoT in agriculture specially for country such as India is very much beneficent and will ensure us in maximum production and less losses for a country where agriculture comprises of the maximum part in the GDP. The farmers can use techniques which checks for soil Ph level, level of water in soil, thus regulating the water supply, level of pesticides, keeping a regulatory check for throughout process form sowing to harvesting a crop. For prediction analysis, we generally use CNN model for analysis; YOLO and OPEN CV can also be used. Using contrast identifying methods or Scada system.

In Surveillance: Nowadays, most of the public areas and buildings are equipped with cameras which are connected with Internet for security purpose for theft, or some misfortunes. All these can be controlled using a mobile or systems also; AI-based motion alert systems or some other things can help us in identifying some issue before hand and thus helping in getting a quicker solution. For this mainly CNN and GPS, GSM-based detection technology is used which predicts such as fire breakouts, and sending an automated message on mobile with location.

In Transportation: The inclusion of IoT in traffic management system especially for centralized traffic monitoring system can be very useful; this can help us in predicting the traffic beforehand and help us to alert the passengers before about the traffic flow and civic authorities to take action based upon the real-time data they are getting through sensors and cameras for easy flow of traffic and vehicles. Also, segregating lanes based on IoT and predictive analysis can help a lot in regulating traffic. Apart from this, we can also use for parking lots for automated parking systems, thus ensuring less traffic on roads and payment systems based on RFID or NFC technology especially, in countries in India where traffic is worst on roads and management is poor also in the times of COVID-19.

These wireless transactions will ensure social distancing. Also, YOLOv3, LSTM, RNN and Mask R-CNN can be used in solving such issues.

6 Comparative Analysis

In the above, Table 1 we can see a complete descriptive point wise analysis on various aspects of the papers that have taken into consideration for this study with their limitations mentioned.

Table 1 Comparison of various paper in this study

Author	Journal name	Procedure	Field	Use cases	Main Purpose	Limitation
Atitallah et al. [23]	ELSEVIER	Detailed study	IoT and smart city	(1) Intelligent applications in home, health care, agriculture, vehicle, surveillance (2) Dl models and applications	Brief study to show an interrelation of smart cities with DL and tech stack used in it	The paper does not involve some new methods such as YOLO, image AI, open cv for smart city analysis adapted in DL
Silva et al. [2]	WILEY | HINDAWI	Detailed study	Big data and smart city	1. Car, 2. Waste, 3. Electricity, 4. Water, 5. Gas, 6. Society	Data which is being processed and was analysed the data was streamlined used for making decisions	Did not included the wide aspects of IoT and their applications with technology used
Mohammadi et al. [3]	IEEE journals	Detailed study	IoT	(a) In home (b) In megalopolis (c) In energy management (d) In Healthline (e) In agriculture (f) In education (g) In sports (h) In retail (i) In IoT-based buildings	For IoT applications and facilities, the most up-to-date deep learning (DL) methods are tested	(1) Does not go into great detail about smart IoT-based city facilities their wide aspects (2) Does not discuss various technical/business problems and concerns which are associated with the growth of IoT-based cities

(continued)

Table 1 (continued)

Author	Journal name	Procedure	Field	Use cases	Main Purpose	Limitation
Mahdavine et al. [4]	DCN	Detailed study	IoT-based megalopolis	(a) In energy (b) In mobility (c) In between citizens (d) In Urban city plan and modernization	Analyses machine learning methods and implementations in IoT-based cities city scenarios	(1) Four cases for IoT-based city implementations are considered: energy, mobility/movement, residents, and urban city planning (2) Does not look at the usage of dl methods for analysing IoT-based city data (3) Does not address the open problems surrounding IoT-based city growth
Zhang et al. [5]	IEEE journals	Detailed study	Cell phone communication networks	–	He studied deep learning methodologies and varied usage over cell phone networks	Does not look into the use of deep learning in various IoT applicability, particularly in the growth of IoT-based cities

(continued)

Table 1 (continued)

Author	Journal name	Procedure	Field	Use cases	Main Purpose	Limitation
Zhang et al. [6]	Information fusion	Detailed survey	Big data analytics	1. Massive quantities of information 2. Data that is heterogeneous 3. Data in real time 4. Data of poor quality	Several studies that use deep learning techniques for big data feature learning is discussed	(1) Just looks at research that use deep learning models for big data function learning.—Does not look into the use of deep learning in various IoT aspects and applicability (2) Does not examine IoT frameworks for smart cities that uses Deep Learning models
Qolomany et al. [7]	IEEE journals	Detailed survey	IoT-based infrastructure	1. Home treatment for the elderly 2. Efficient energy use 3. Convenience and amusement 4. Protection and safety 5. Programmes that are not related to the others	Review the usage of ML and BDA for the growth of smart infrastructure	(1) Does not concentrate in IoT-based city applications (2) Does not go over DL methods (3) Does not examine the use of machine learning models in fog/edge computing

(continued)

Table 1 (continued)

Author	Journal name	Procedure	Field	Use cases	Main Purpose	Limitation
Chen et al. [8]	IEEE journals	Detailed survey	IoT-based city	1. Getting around 2. Medical treatment 3. Public safety 4. Environment	It looks for a detailed study on DL in IoT-based cities, summarizes the most common models, the most recent work on DL-based smart city applications	(1) Excludes IoT-based home and agriculture applications (2) Does not look at how cloud, fog, and edge computing can be used for DL-based applications (3) Does not address the issue of IoT big data

7 Challenges

The biggest challenge or issue that the smart cities nowadays are facing is security and privacy issue. The people who are having malicious intent exploit the databases and get all this raw data and sell this into the dark web or ask for huge ransoms from the company to keep their data private. In the past year, there have been many such incidents such as in New York; a hacker hacked around 200 homes security camera and leaking that data in deep webs; the deep web comprises the 90% or more part than what we access in daily life. It was developed by US for top secret and confidential activities to take place. Later it became something which was completely opposite of it. Recently, in March hackers breached a company Verkada Inc., all surveillance cameras getting an access to 150 k cameras thus ruining up the reputation of company for their own profit.

Previous year, there was an attack of Ransomware which actually affected many organizations and their data, causing them huge losses. Some notable companies are which were affected by it:

- Cognizant
- Magellan Health
- CPI—Communication and Power Industries
- University of California San Francisco
- Baltimore County Public Schools
- Advantech
- Carnival Corporations
- Canon.

All these companies since then have implemented two factor authentications for account safety and protecting their databases. Also, many incidents happened with companies like dominos all the user base data got leaked with all credit card details of customers being exposed in dark web. Apart from this, there have been many such cases reported all such activities which have taken place in the past are the limitations of IoT and security they have.

The attacks usually used by the hackers to gain access to these IoT-based systems are:

- Botnets Attack
- Denial of Service Attack
- Man in the Middle Attack
- Identify and Data Theft
- Social Engineering
- Advanced persistent threats
- Ransomwares
- Remote Recording.

Also, apart from the security issues there are some other issues which are related to the Hadoop and Big data which is the management of nodes and high-end infrastructure everywhere is not possible specially country like India which is still developing where advancement is supposed to go step by step not in haste. Not everyone can buy the required devices and build the infrastructure. As, if there is some disaster it requires racks and mostly companies in India rely on single racks which is very wrong as there is no disaster management for databases. In India, not large datasets are available; thus, we are not able to get the very accurate data to make predictions as the foreign countries they use it for analysis and making decisions. Cost of production and implementation for IoT-based systems is too high apart from this if we implement, we need someone to monitor and make it function thus, mostly companies do not want to put that much effort until the work gets done. Also, the quality of service in India is not up to the mark as no production plants are there who manufacture on a large base; thus, we need to import everything from outside country which is increasing the inflation rate and cost of the good.

Through the above study, we saw there are numerous challenges consolidated altogether here in a summarized manner which is: some papers do not focus on the newer research topics and technology and some are defined to a particular technology whether it is a smart home, infrastructure, or big data processing data in real time. Also, some papers are based on DL models. The comparison made on different areas of the smart city is confined to only some areas not covering the wider spectrum. As IoT is a wide field and has expanded a lot in the recent trends since 2007 and still its demand is growing day by day and so is the concern for the data privacy.

8 Conclusion

We can conclude that our paper was segregated into the basic introduction which lets us know about basic IoT and how it is useful to us. Then, we showed some related works done on these smart IoT-based cities, after that we showed a visual analysis of different recent trends that took place and the variation between them creating demand for big data with the planned estimated investment for the same for providing better applications and life experience. After this, we talked about various other technologies which help it to function properly. Thereafter, we talked about the methodologies used in these cities and how the recent trends are happening on this area with the advent of pandemic. This led us to the observation with the related surveys we did and we were successfully able to show the survey and detailed analysis basically a juxtaposition on different aspects of the research work done based on the parameters in the direction of "Technologies Involved in IoT-based Smart Cities". Followed by, the challenges faced in such cities which covered all recent happenings around the world along with some technological and financial crisis which will help us to understand this on a better scale.

References

1. What is IoT and some predictions based on Oracle's documentation https://www.oracle.com/in/internet-of-things/what-is-iot/
2. Silva BN, Khan M, Han K, Big data analytics embedded smart city architecture for performance enhancement through real-time data processing and decision-making. https://doi.org/10.1155/2017/9429676
3. Mohammadi M, Al-Fuqaha A, Sorour S, Guizani M (2018) Deep learning for IoT big data and streaming analytics: a survey. IEEE Commun Surv Tutor 20:2923–2960. http://arxiv.org/abs/1712.04301
4. Mahdavinejad MS, Rezvan M, Barekatain M, Adibi P, Barnaghi P, Sheth AP (2018) Machine learning for internet of things data analysis: a survey. Digit Commun Netw 4:161–175. DCN https://www.sciencedirect.com/science/article/pii/S235286481730247X
5. Zhang C, Patras P, Haddadi H (2018) Deep learning in mobile and wireless networking: a survey. IEEE Commun Surv Tutor. http://arxiv.org/abs/1803.04311
6. Zhang Q, Yang LT, Chen Z, Li P (2018) A survey on deep learning for big data. Inf Fusion 42:146–157. https://doi.org/10.1016/j.inffus.2017.10.006
7. Qolomany B, Al-Fuqaha A, Gupta A, Benhaddou D, Alwajidi S, Qadir J, Fong AC (2019) Leveraging machine learning and big data for smart buildings: a comprehensive survey. IEEE Access 7:90316–90356. https://ieeexplore.ieee.org/abstract/document/8754678/
8. Chen Q, Wang W, Wu F, De S, Wang R, Zhang B, Huang X (2019) A survey on an emerging area: deep learning for smart city data. IEEE Trans Emerg Top Comput Intell 3:392–410. https://ieeexplore.ieee.org/abstract/document/8704334/
9. Begam SS, Selvachandran G, Ngan TT, Sharma R (2020) Similarity measure of lattice ordered multi-fuzzy soft sets based on set theoretic approach and its application in decision making. Mathematics 8:1255
10. Vo T, Sharma R, Kumar R, Son LH, Pham BT, Tien BD, Priyadarshini I, Sarkar M, Le T (1 Jan 2020) Crime rate detection using social media of different crime locations and twitter part-of-speech tagger with brown clustering. 4287–4299
11. Nguyen PT, Ha DH, Avand M, Jaafari A, Nguyen HD, Al-Ansari N, Van Phong T, Sharma R, Kumar R, Le HV, Ho LS, Prakash I, Pham BT (2020) Soft computing ensemble models based on logistic regression for groundwater potential mapping. Appl Sci 10:2469
12. Jha S et al (2019) Deep learning approach for software maintainability metrics prediction. IEEE Access 7:61840–61855
13. Sharma R, Kumar R, Sharma DK, Son LH, Priyadarshini I, Pham BT, Tien Bui D, Rai S (2019) Inferring air pollution from air quality index by different geographical areas: case study in India. Air Qual Atmos Health 12:1347–1357
14. Sharma R, Kumar R, Singh PK, Raboaca MS, Felseghi R-A (2020) A systematic study on the analysis of the emission of CO, CO_2 and HC for four-wheelers and its impact on the sustainable ecosystem. Sustainability 12:6707
15. Sharma S, et al (29 Oct 2020) Global forecasting confirmed and fatal cases of COVID-19 outbreak using autoregressive integrated moving average model. Front Public Health https://doi.org/10.3389/fpubh.2020.580327
16. Malik P, et al (15 Jan 2021) Industrial internet of things and its applications in industry 4.0: state-of the art. Comput Commun 166:125–139. Elsevier
17. Analysis of water pollution using different physico-chemical parameters: a study of Yamuna river. (11 Dec 2020) Front Environ Sci. https://doi.org/10.3389/fenvs.2020.581591
18. Dansana D, Kumar R, Parida A, Sharma R, Adhikari JD et al (2021) Using susceptible-exposed-infectious-recovered model to forecast coronavirus outbreak. Comput, Mater Continua 67(2):1595–1612
19. Vo MT, Vo AH, Nguyen T, Sharma R, Le T (2021) Dealing with the class imbalance problem in the detection of fake job descriptions. Comput, Mater Continua 68(1):521–535
20. Sachan S, Sharma R, Sehgal A (2021) Energy efficient scheme for better connectivity in sustainable mobile wireless sensor networks. Sustain Comput: Inf Syst 30:100504

21. Ghanem S, Kanungo P, Panda G, et al (2021) Lane detection under artificial colored light in tunnels and on highways: an IoT-based framework for smart city infrastructure. Complex Intell Syst https://doi.org/10.1007/s40747-021-00381-2
22. Sachan S, Sharma R, Sehgal A (2021) SINR based energy optimization schemes for 5G vehicular sensor networks. Wireless Pers Commun. https://doi.org/10.1007/s11277-021-08561-6
23. Atitallah SB, Driss M, Boulila W, Ghézala HB, Leveraging deep learning and IoT big data analytics to support the smart cities development: review and future directions. https://doi.org/10.1016/j.cosrev.2020.100303

Vibration Analysis of Fluid Structure Interface for Rectangular Plate

Kayser Aziz Ameen, Husam Abdulrasool Hasan, Jenan S. Sherza, Hiba A. Hasan, Raheem J. Mohy, and Ali A. Hatam

Abstract Investigate the vibration of the rectangular plate related to the liquid, normal body frequency and modes shapes of cantilever plate without and with hole and perforate plate with 169 hole in air and contact of the water surface and immersed in water are presented using finite element method via ANSYS15 software. Acoustic model in three dimension domain is considered using APDL program to take the variables. The gotten comes about detailed the 6th plate normal body frequency and mode shapes which are based upon the behavior of the plate. In all cases, there are diminished with in the natural frequency of the fluid–structure framework. It can be concluded from our work that the exactness of the predicted frequencies utilizing our demonstrate is either exceptionally great or at slightest adequate for commonsense purposes.

Keywords Vibration · Natural frequency · Mode shape · ANSYS · Finite element method · Open tank · Acoustic analysis · Modal analysis · Fluid–structure interface · Coupled problem

---

K. Aziz Ameen (✉) · H. Abdulrasool Hasan · J. S. Sherza · R. J. Mohy · A. A. Hatam
Al-Esraa University College, Air Conditioning and Refrigeration Department, Baghdad, Iraq
e-mail: dr.qaysar@esraa.edu.iq

H. Abdulrasool Hasan
e-mail: drhusam@esraa.edu.iq

J. S. Sherza
e-mail: jenansherza@esraa.edu.iq

A. A. Hatam
e-mail: ali.hatem@esraa.edu.iq

H. A. Hasan
Computer Techniques Engineering Department, Al-Turath University College, Baghdad, Mansour, Iraq
e-mail: hiba.abdalameer@turath.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_4

1 The Preface

The issue of vibrating plates in contact with Liquid has interested numerous examiners, either in the past or in the present. This issue may be found in different sorts of designing areas: such as holders, frames beneath or mostly submerged in liquid such as water, dams and conduits, course cloughes, the frame, furthermore, limited vibrations of boats and submarine, etc. It is by and large realized that the eigen frequencies of constructions in touch with water diminish essentially contrasted with those in void, particularly for the elemental “eigen frequency.” It is notable that the characteristic frequencies of structures in contact with liquid are distinctive of what is under discussion. In this manner, the expectation of common recurrence “changes due to” the nearness of the liquid is critical for planning the frames in contact with or drenched in liquid. Marcus [1] examined the vertically free vibration and even immerse the cantilever slab an interminable liquid space by the limited component strategy.

Bauer [2] considered the association of an essentially upheld flexible of rectangular holder having inflexible divider and filled with fluid, there is exact free zone solution. As of late, Soedel and Soedel [3] examined the free and constrained vibration of a basically bolstered tablet associated to a huge store of fluid on the four sides and support a fluid with openly sloshing surface. Kwak [4] considered the non-dimensional included virtual mass gradual components of a perplexed rectangular slab in touch with liquids like water by the joined use of “Rayleigh Ritz technique” and “Green’s work technique.” The impact of liquids such as water profundity on the fiery direct of free edge round plates was examined by “Kwak and Han” [5], and they received the Hankel change strategy, which brought about in double necessarily conditions. The arrangement of double indispensably conditions was determined mathematically utilizing “Fourier–Bessel” arrangement. The “Kirchhoff hypothesis” of plates is utilized to demonstrate the versatile lean plate. “Zhou and Cheung” [6] considered the vibratory attributes of a rectangular plate in touch with liquids such as water on one face. The adaptable plate is viewed as a part of an upward rectangular resolute divider in touch with water, the edges of which are deftly restricted and parallel to those of the unbending divider. The case of a circular plate composing the base of a unbending round and hollow liquid-filled tank was considered by “Cheung and Zhou” [7] counting thought of the liquid free surface impact. Kerboua et al. [8], they created approach suitable for vibration detection of rectangular plate coupled with liquid. This case represented some key components of a complex structure utilized in businesses such as aviation, atomic, and maritime. The slab can be completely immersed in a liquid or glide on their free surface.

In this studied, a greater widespread case of the interplay of a square plate in touch with water is considered: The plate is located in an open square tank of a vertical square strong wall in touch with water. The edges of the panel are parallel to the edges of the wall, and the panel is located in the center of the tank. Such a model can be used for hydro-vibration analysis of flood gates [9–27], sewer manholes, etc.

2 Vibroacoustic Analysis

"The finite element method (FEM)" used for plate analysis in this study. The "finite element method (FEM)" is utilized in this case to behavior plate examination that surrounded by water and study the natural frequencies and mode shapes resulting from vibration plate. The effect of different media (air, on water surface and immersed into the water) and configuration of plate on the vibration is presented. The FEM is consisted of modeling the fluid of the medium and the fluid surrounding the structure (fluid structure interface FSI).

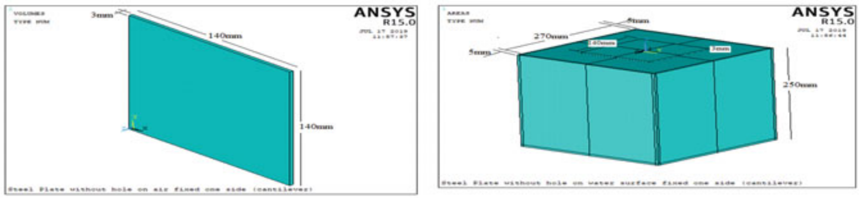
In the analysis, within the sound investigation, modeling is doing for weight in fluid for diverse frequencies, molecule speed, sound weight level, constriction, radiation and scattering of sound waves, etc. Within the investigation, the proper analysis was taken into account the interaction of the coupling structure with the liquid. The FEM is imposed the liquid was incompressible, but only a relatively minor change of pressure was allowing depending on the average pressure. The FEM is imposes the viscous substance present in the fluid nothing to do with the dissipation of acoustic energy. The FEM is enforced both the mean (pressure, density) and it imposes the pressure that resolves, it is a deviation from the relative pressure and not the absolute pressure [9].

3 Geometrical and Material Properties of the Fluid–Structure

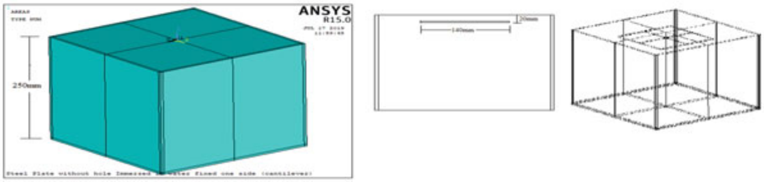
In the current study, three models are designed. The first case is the plate without hole, treated with three cases in air, contact the water surface, and immersed in water as shown in Fig. 1. Illustrated (open tank–up side) (outer dim. = 280 × 280 × 255 mm^3, inner dim. = 270 × 270 × 250 mm^3) and volume of water = 270 × 270 × 250 mm^3.

The boundary condition is fixed one end—free other three ends (i.e., cantilever plate) and the tank is open tank (from the upper)—fixed on its base. The materials: plate**:** steel—"E = 200 GPa, $\nu = 0.3$, $\rho = 7850 Kg/m^3$" and the tank**:** steel—"E = 200 GPa, $\nu = 0.3$, $\rho = 7850 Kg/m^3$." Too it can be watched from picture (1c) that the plate immersed in the water tank (open tank–upper side) with depth = 20 mm. The second case is the plate with central hole, and two cases are tested one for $R =$ 20 mm and the other with $R = 10$ mm. Also treated with three cases in air, contact the water surface and immersed in water as shown in Fig. 2.

The third case is perforated plate with 169 holes as shown in Fig. 3.

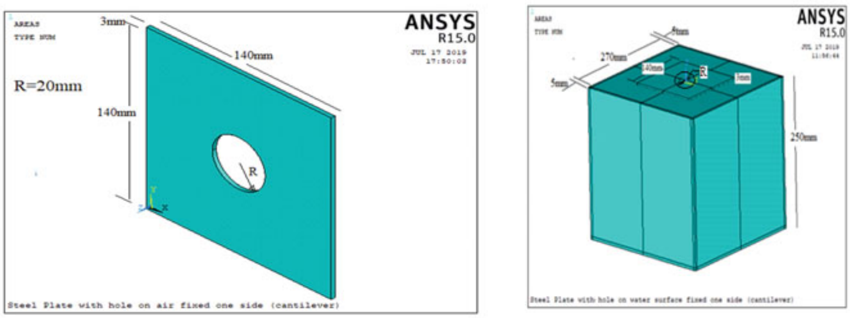


(a) Plate Dim.140x140x3 mm (b) Plate on the water surface Tank



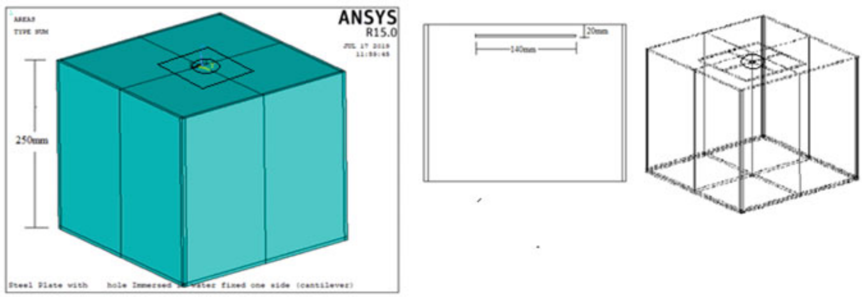
(c) Plate immersed in water

Fig. 1 Plate without hole **a** on air, **b** on water surface, **c** immersed in water



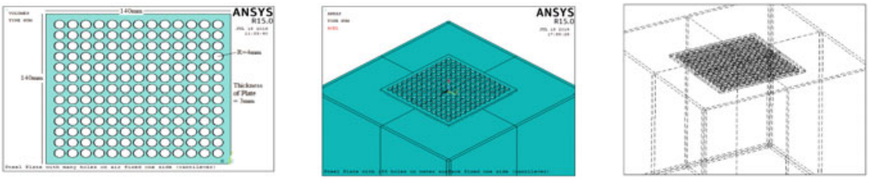
(a) Plate Dimensions
Radius of the hole(R)=20 and R=10mm

(b) Plate with hole on the water surface Tank



(c) Plate immersed in water

Fig. 2 Plate with central hole **a** on air, **b** on water surface, **c** immersed in water



(a) Plate Dim., Radius of the water surface Tank
No. of holes=13x13=169

(b) Plate with holes on immersed in water

(c) Plate with holes the hole(R)=4 mm

Fig. 3 Perforated plate with 169 hole **a** on air, **b** on water surface, **c** immersed in water

4 ANSYS Model

Coupled fluid–structure treated in this paper as acoustic model. In acoustic elements which are contacted with the structures (FLUID 29 and FLIDID 30), it uses (KEYOPT (2) = 0) for assumption action which is that the structure interacts with fluids, thus leading to use of element matrices, unsymmetric with degrees of freedom (UX, UY, UZ). KEYOPT (2) = 1, it is used for all remaining acoustic elements, where fluid is putted inside the structure. In this case, the symmetric matrix is used, and it is less cost in terms of time and storage. Therefore, it is best to use them if the situation permits. Figure 4 shows the interaction between structure and fluid [10].

Auxiliary acoustic examination the plates with water and without both can be implemented in “ANSYS” utilizing “ANSYS/Multiphysics and ANSYS/Mechanical programs.” The acoustic examination in “ANSYS” is for the most part done by modeling structure space, modeling liquid space depending on outside or inner issue with fitting liquid components. Acoustic shows by and large comprises of basic space liquid space and unbounded acoustic space. “ANSYS” employments components “FLUID129 for 2D and FLUID130” for 3D which are the boundless acoustic components whereas “FLUID29 for 2D and FLUID30” for 3D are limited acoustic

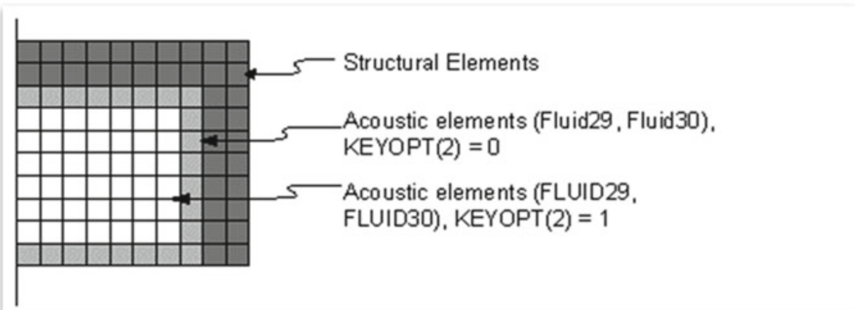


Fig. 4 2D acoustic model

components. Boundless acoustic component modeled as a circle in 2D and circle in 3D. It assimilates all weight waves created by a source without any reflections.

Unlimited elements cannot be used with basic elements. The elements are used to model the structure and acoustic space. Solid elements are used for lean plate or shell structures of appropriate thickness. The “solid 185” component is used for structural modeling. There are eight nodes, and each node has three degrees of freedom interpretation.

“FLUID30” is used to simulate fluid elements. It is a 3D element with 4 nodes. It is used for coupled and uncoupled fluid elements. Coupling elements have three degrees of freedom “Ux, Uy, and pressure,” while non-coupling elements have only one. The simulation model is composed of structure domain, fluid domain, and fluid–solid interface. The structure domain is modeled by solid 185 elements, and the fluid domain is modeled by “FLUID30.” Specify the flow-structure interface on the common node of the structure and the fluid domain. Figure 5 shows the “ANSYS” mesh for all cases.

“FLUID30” is utilized to demonstrate fluid element. It can be a 3D detail which has four nodes. It changed into applied for each coupled and uncoupled liquid elements.

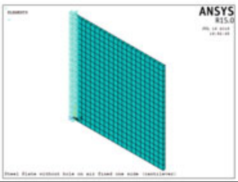
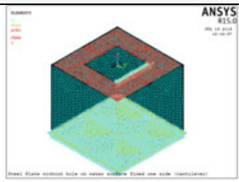
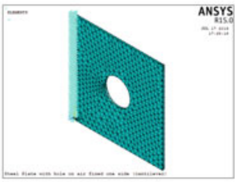
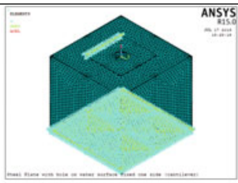
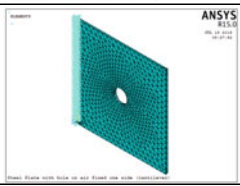
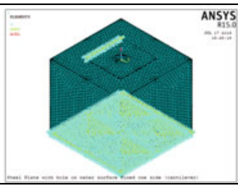
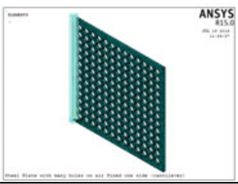
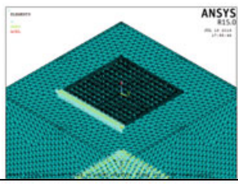
Mode No.	in the air	On the water surface
Model		
Model		
Model		
Model		

Fig. 5 ANSYS mesh

Coupled detail has 3 levels of freedom "UX, UY, and pressure," while an uncoupled detail has as it were one. Reenactment show comprises of structure space, liquid space, and fluid–structure interface. Structure space is demonstrated by utilizing solid 185 element, whereas liquid space is demonstrated by utilizing "FLUID30." Fluid-frames interface becomes indicated on not unusual place nodes for frame and liquid space. Figure 5 appears the "ANSYS" work for all cases.

5 Modal Analysis Procedure

Modal analysis is applied to solution all the dynamic problems of structures, and to know vibration characteristics of structures, such as (natural frequencies and modes).

There are connected four steps when utilizing modular examination [11]:

- Create the model
- Choose examination sort and options.
- Put the degree of freedom which known on the final $[K]$ matrix.
- Results overview.

The essential condition illuminated in a normal modular examination is the classical eigenvalue problem. The physical frequencies and mode shapes of a shape are quite important parameters within the plan of a structure for energetic stacking conditions [12]. The un-damped essential condition illuminated by the modular examination is

$$[k]\{w\} + [M]\{\ddot{w}\} = \{0\} \tag{1}$$

Assuming harmonic motion that is

$$\{w_i\} = \{\varphi_i\} \sin \omega_i t; \quad i = 1, 2, \ldots n \tag{2}$$

where n is the number of variables in the system to be found $\{\varphi_i\}$ = The mode form vector for the ith mode of vibration, and ω_i = the "angular frequency" of mode i.

We drive equation no. (2) with regard to time yields

$$\{\ddot{w}_i\} = -\omega_i^2 \{\varphi_i\} \sin \omega_i t \tag{3}$$

At that point, substituting conditions (2) and (3) into condition (1) yields, after erasing the expression "term ($\sin \omega_i$):"

$$[K] - \omega_i^2 [M]\{\varphi_i\}\{\varphi_i\}^T = \{0\} \tag{4}$$

Condition (4) has the form of the arithmetical eigenvalue problem "($K\phi = \lambda M\phi$)." From the speculation of "homogenous equations," non-trivial preparations exist because it had been in case the determinant of the coefficient system is wreck despite to zero. Development of the determinant yields a polynomial of set up NR known as feature equation.

The NR roots of this polynomial $\left(\omega_i^2\right)$ are the feature values or the eigenvalues. Substitution of those roots (one at every time) into the homogeneous condition (4) generated the feature vectors or the eigenvectors $\{\phi_i\}$ inside subjective constants. On the alternative hand, every eigenvector can be determined as any column of borders arrange $\left[H_i^a\right]$ of the feature framework [Hi], gotten from condition (4) as take after:

$$[H_i]\{\varphi\} = [0] \tag{5}$$

where

$$[K] - \omega_i^2[M] = [H_i] \tag{6}$$

Subspace approach is employed as a numerical strategy in this research to solve the upward condition using "FEM" [13].

$$[K] - \omega_i^2[M] = \{0\} \tag{7}$$

6 Results and Discussion

From this field that the "natural frequencies" of frames which a touch any liquid are diverse from other influential surrounding. Subsequently, the forecast of "natural frequency" variation because of the closeness of the liquid is imperative for planning frames which are in contact with or inundated in liquid.

In common, the impact of the liquid constrain on the structure is spoken to as included mass, which brings down the common frequency of the shape from that which could be measured in the air.

This diminish within the normal recurrence of the fluid–structure framework is caused by expanding the active energy of the coupled framework without a comparing increment in strain energy.

An exceptional sum of exertion has been carried out on issues including energetic interaction between an flexible structure and a encompassing liquid medium.

The variation of fluid level is considered as a constant in the calculation of the natural frequencies and its mode shapes. Figure 6 illustrated the distribution of the natural frequencies for all cases. Figures 7, 8, 9, and 10 illustrated the mode shapes for all cases.

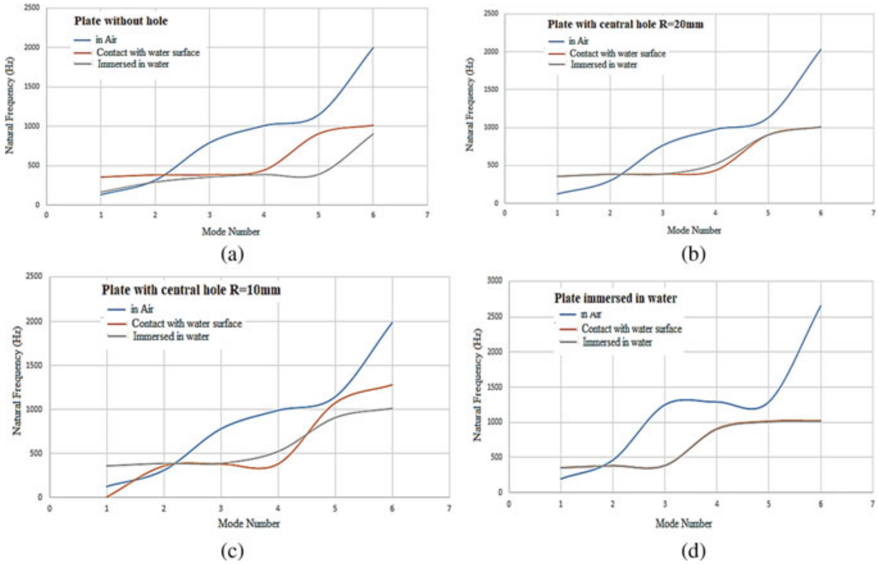


Fig. 6 Natural frequencies **a** plate without hole, **b** plate with central hole $R = 20$ mm, **c** plate with central hole $R = 10$, **d** plate with perforated plate (169 holes)

7 Inference

- The reason of the examination portrayed in this study is to decide common frequencies of square slab under water, which constitute critical auxiliary components in different segments of manufacture. The slabs may be totally immersed or drifting on the free flatness of fluid. Normal recurrence is influenced by the included mass related with the inactivity of the liquid that is constrained to waver when the structure vibrates.
- In all cases, there are diminished within the "natural frequency" of the fluid–structure framework.
- It may be concluded from our paintings that the exactness of the frequencies calculated making use of our exhibit is both extraordinarily top or at slightest ok for common sense purposes.
- It may be proven that the normal frequency of the plate extended with growing the mode number.
- It may be seemed that the normal frequency of the slab increased with increasing the mode number.

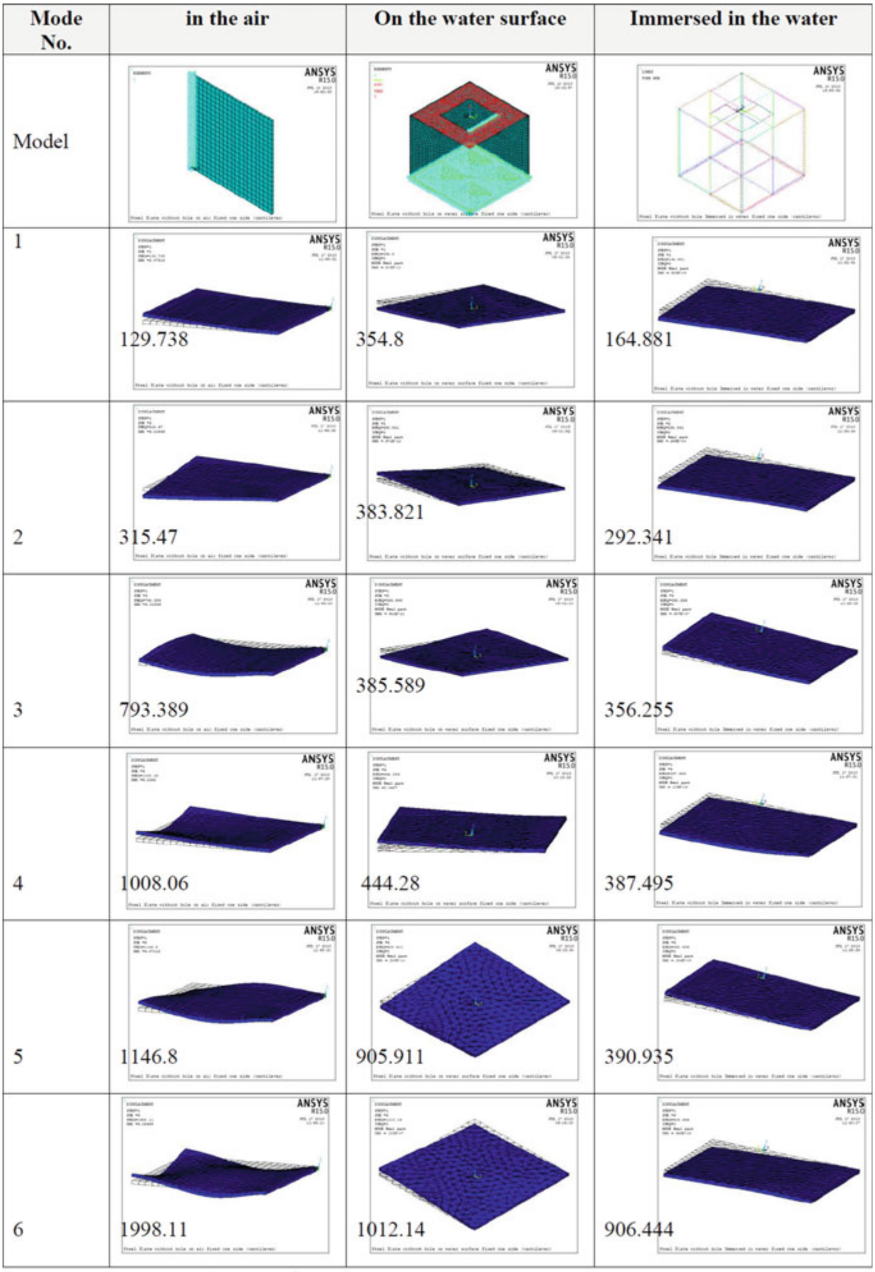


Fig. 7 Case-1-"plate without hole" natural frequencies and mode shapes

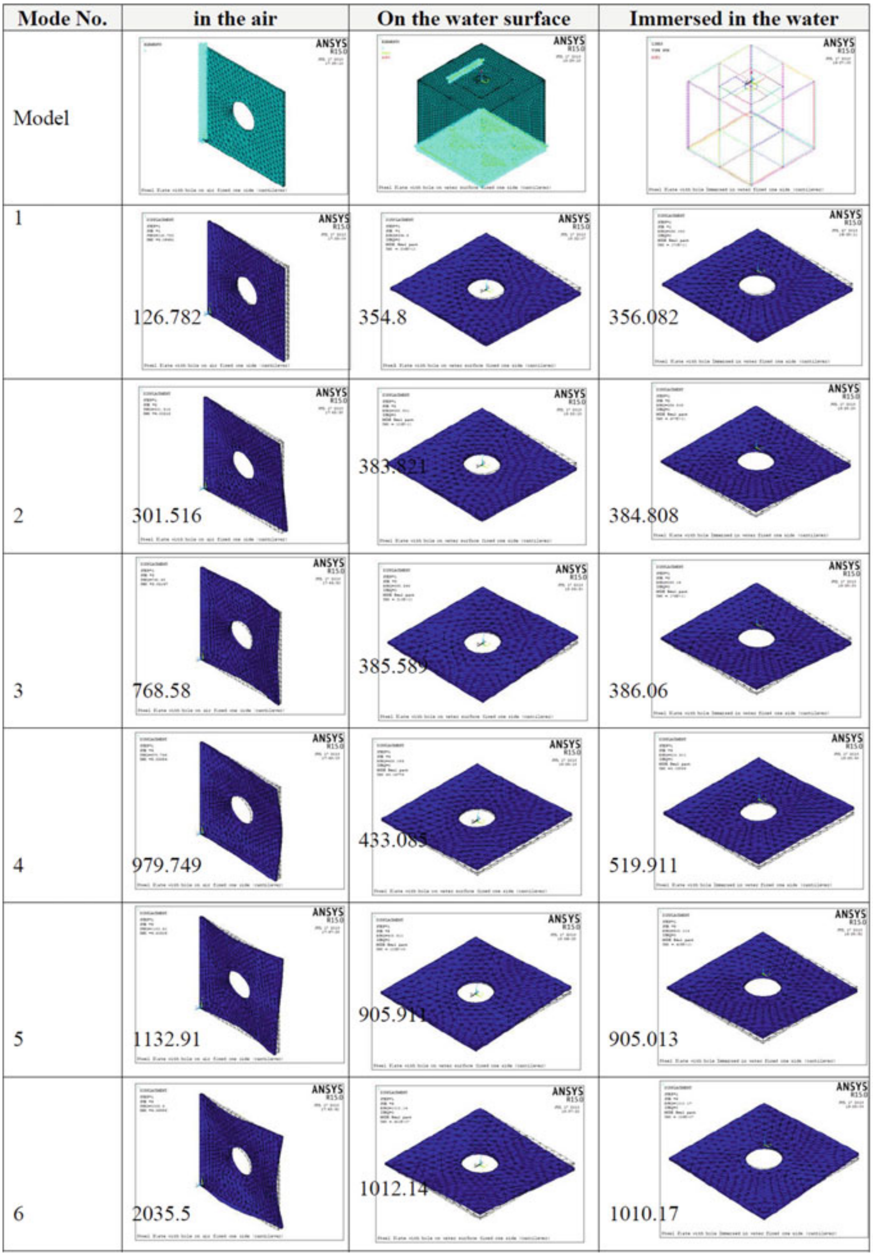


Fig. 8 Case-2- I-plate with central hole—natural frequencies and mode shapes for $R = 20$ mm

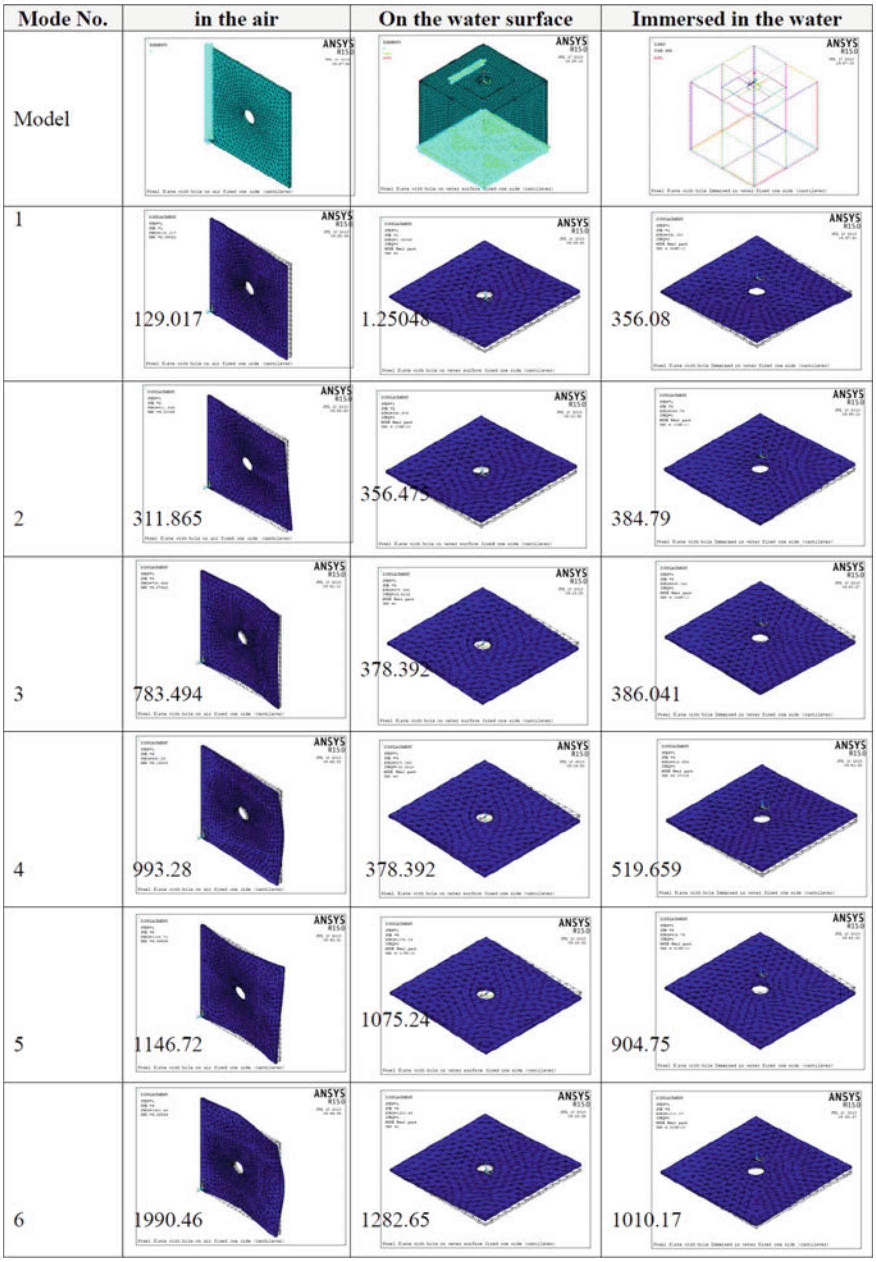


Fig. 9 Case-2-II-plate with central hole—natural frequencies and mode shapes for $R = 10$ mm

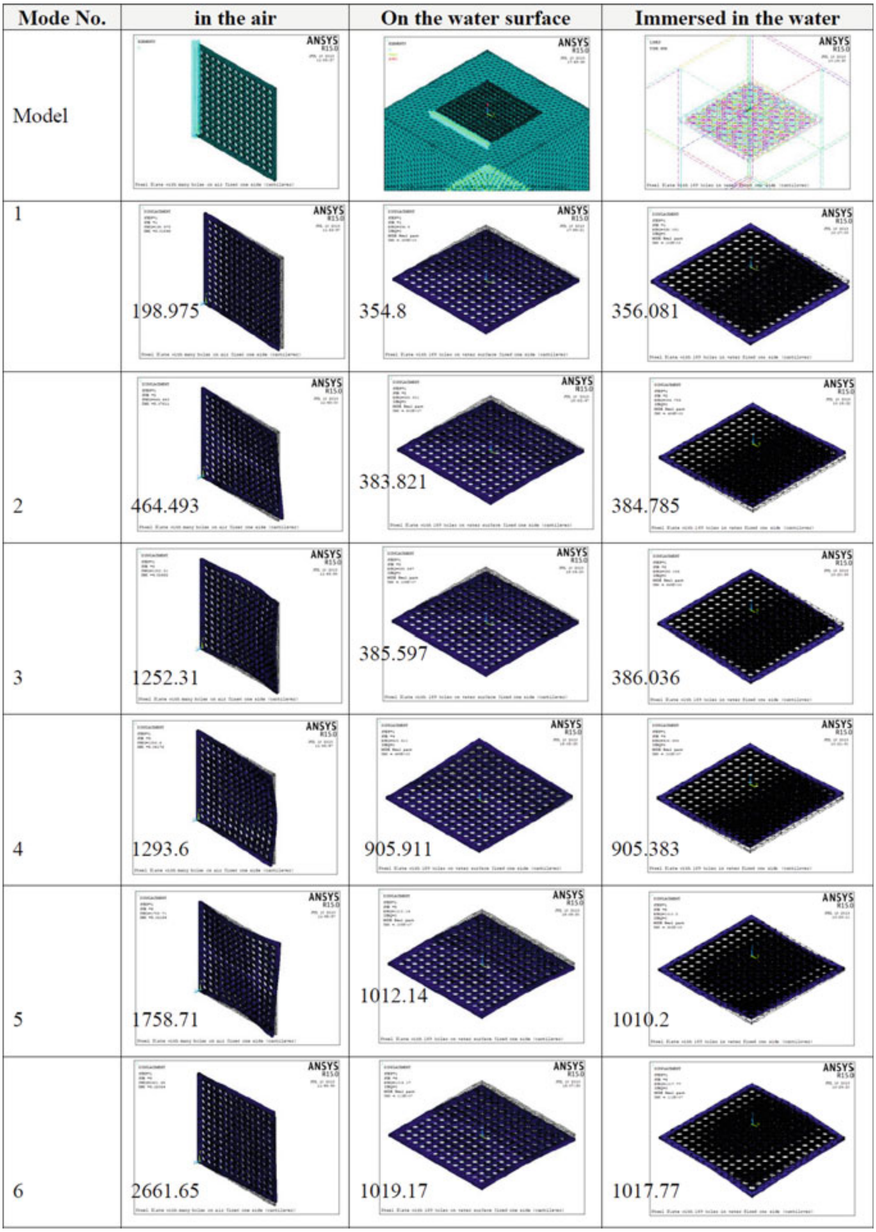


Fig. 10 Case-3-perforate plate—natural frequencies and mode shapes for plate with 169 holes $R = 4$ mm

References

1. Marcus MS (1978) A finite-element method applied to the vibration of submerged plates. J Ship Res 22:94–99
2. Bauer HF (1981) Hydroelastic vibrations in a rectangular container. Int J Solids Struct 17:639–652
3. Soedel SM, Soedel W (1994) On the free and forced vibration of a plate supporting a freely sloshing surface liquid. J Sound Vib 171:159–171
4. Kwak MK (1996) Hydroelastic vibrations of rectangular plates. J Appl Mech ASME 63:110–115
5. Kwak MK, Han SB (2000) Effect of fluid depth on the hydroelastic vibration of free-edge circular plate. J Sound Vibrat 230(1):171–185
6. Zhou D, Cheung YK (2000) Vibration of vertical rectangular plate in contact with water on hone side. Earthq Eng Struct Dyn 29:693–710
7. Cheung YK, Zhou D (2002) Hydroelastic vibration of a circular container bottom plate using the Galerkin method. J Fluid Struct 16(4):561–580
8. Kerboua Y, Lakis AA, Thomas M, Marcouiller L (2008) Vibration analysis of rectangular plates coupled with fluid. Appl Math Model 32:2570–2586
9. Fahy FJ (2012) Sound and structural vibration: radiation, transmission and response 7.2, numerical analysis of sound fields. Technology and engineering, pp 273–274. Coupled field, acoustics (UP19980818). www.ansys.stuba.sk/html/guide_55/g-cou/GCOU4.htm.ANSYS Basic analysis procedures guide. http://www.ansys.stuba.sk/html/guide_55/g-bas/GBASToc.htm
10. Ameen HA, Ismail AH, Mashloosh KM (2012) Effect of stiffeners and thickness of shell on the natural frequencies and mode shapes of oblate shell by finite element method. Al-Qadisiya J Eng Sci 5(3):325–340; ANSYS, Release (15.0) (2014) ANSYS mechanical APDL element reference and ANSYS mechanical APDL structural analysis guide. ANSYS, Inc., USA
11. Gadicha AB et al (2021) J Phys: Conf Ser 1963:012141
12. Hasan HA, Sopian K, Fudholi A, Photovoltaic thermal solar water collector designed with a jet collision system. Energy 161. j.energy.2018.07.141
13. Naje AS, Chelliapan S, Zakaria Z, Ajeel MA, Sopian K, Hasan HA (2016) Electrocoagulation by solar energy feed for textile wastewater treatment including mechanism and hydrogen production using a novel reactor design with a rotating anode. RSC Adv 6(12):10192–10204
14. Aboghrara AM, Baharudin BTHT, Alghoul MA, Adam NM, Hairuddin AA, Hasan HA (2017) Performance analysis of solar air heater with jet impingement on corrugated absorber plate. Case Stud Therm Eng 10:111–120. https://doi.org/10.1016/j.csite.2017.04.002
15. Hameed Jaaz A, Hasan HA, Sopian K, Kadhum AAH, Gaaz TS, Al-Amiery AA (2017) Outdoor performance analysis of a photovoltaic thermal (PVT) collector with jet impingement and compound parabolic concentrator (CPC). Materials (Basel) 10(8). https://doi.org/10.3390/ma10080888
16. Sopian K, Alwaeli AHA, Hasan HA, Al-Shamani AN (2018) Advances in high efficiency photovoltaic thermal solar collectors. J Adv Res Fluid Mech Therm Sci 47(1)
17. Hasan HA et al (2017) Numerical investigation of Microjet impingement of water for cooling photovoltaic solar cell. J Adv Res Fluid Mech Therm Sci 51(1):1108–1121. 2017.icheatmasstransfer.2012.10.023
18. Al-Shamani AN, Sopian K, Mat S, Hasan HA, Abed AM, Ruslan MH (2016) Experimental studies of rectangular tube absorber photovoltaic thermal collector with various types of nanofluids under the tropical climate conditions. Energy Convers Manag 124:528–542. https://doi.org/10.1016/j.enconman.2016.07.052
19. Sopian K et al (2017) Effect of the nozzle exit position on the efficiency of ejector cooling system using R134A. ARPN J Eng Appl Sci 12(18):5245–5250
20. Hasan HA, Alquziweeni Z, Sopian K (2018) Heat transfer enhancement using nanofluids for cooling a central processing unit (CPU) system. J Adv Res Fluid Mech Therm Sci 51(2)

21. Ameen KA, Hasan HA, Al-Dulaimi MJ, Abed AM, Al-Qrimli HF (Jan 2022) Improving the performance of air conditioning unit by using a hybrid technique. MethodsX 9:101620. https://doi.org/10.1016/j.mex.2022.101620
22. Hasan HA, Sopian K, Jaaz AH, Al-Shamani AN, "Experimental investigation of jet array nanofluids impingement in photovoltaic/thermal collector," Sol. Energy, vol. 144, pp. 321–334, 2017.solener.2017.01.036.
23. Mohammed HA, Hasan HA, Wahid MA (2013) Heat transfer enhancement of nanofluids in a double pipe heat exchanger with louvered strip inserts. Int Commun Heat Mass Transf 40(1):36–46
24. Rukman NS, et al (Jul 2021) Bi-fluid cooling effect on electrical characteristics of flexible photovoltaic panel. J Mechatron, Electr Power, Veh Technol 12(1):51–56. 2021.v12.51–56
25. Sopian K et al (2017) A stand-alone photovoltaic system design and sizing: a greenhouse application in Sabha city: case study in Libya. Proc 3rd Eng Sci Technol 3:1–10
26. Dezfouli MMS et al (2017) Energy saving potential of solar cooling systems in hot and humid region. ARPN J Eng Appl Sci 12(18):5241–5244
27. Ameen KA, Al-Dulaimi MJ, Abrahem HA, Hasan HA (2020) Experimental study to increase the strength of the adhesive bond by increasing the surface area arrangement. IOP Conf Ser Mater Sci Eng 765:1–10

IoT-Based Prediction of Chronic Kidney Disease Using Python and R Based on Machine and Deep Learning Algorithms

V. Shanmugarajeshwari and M. Ilayaraja

Abstract The machine learning (ML) and Internet of things (IoT) technologies are increasingly focussed on decision tree classification algorithm. Its use is expanding through numerous fields, solving incredibly complex problems. DTCA is also being used in medical health data using computer-aided diagnosis to identify chronic kidney diseases like cancer and diabetes. Deep learning is a class of machine learning that utilizes neural networks to solve problems and learn unsupervised from unstructured or unlabelled data. The DL used to deep stacked auto-encoder and softmax classifier methods is applied for CKD. In this work, based an R Studio and Python Colab software using random forest, SVM, C5.0, decision tree classification algorithm, C4.5, ANN, neuro-fuzzy systems, classification and clustering, CNN, RNN, MLP is used to predict multiple machine and deep learning techniques, discover an early diagnosis of CKD patients. In this work, classify the chronic kidney disease various stages are identified.

Keywords Decision tree classification algorithm · Computational decision support system · Chronic kidney disease · Deep and machine learning algorithms · IoT and CKD stages

1 Introduction

One of the fastest rising non-communicable diseases is chronic kidney disease or simply CKD, which contributes to a big problem of death and illness. Now 2020, CKD beat 803 million individuals worldwide, with 663 million males and 526 million females [1]. CKD is also a huge public medical health issue in India, and this is related to 17% of the world's population [2]. Analytical environments are employing a variety

V. Shanmugarajeshwari (✉) · M. Ilayaraja
Department of Computer Applications, Kalasalingam Academy of Research and Education, Krishnankoil, Virudhunagar, Tamil Nadu 626126, India
e-mail: v.shanmugarajeshwari@gmail.com

M. Ilayaraja
e-mail: ilayaraja.m@klu.ac.in

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_5

of ways to improve the value of health-related problem prediction by developing and exploring healthcare data records. To analyse various types of documents, a diverse range of procedures is required [3]. The kidneys' job is to filter blood and pass it through a filter. It removes unnecessary blood and maintains electrolyte and hydration balance. It strains blood and produces urine, which is produced by the kidney's two bean-shaped structures. Every kidney is surrounded by a million nephrons units of measurement [4].

a. CKD Risk Factors

The main cause of CKD is diabetes, although additional variables that contribute to the disease include hypertension, obesity, heart disease, family history, alcohol, and age.

b. Symptoms

CKD symptoms of the urinary_functionality alterations, serum in the urine, swelling and soreness, unexpected tiredness, and tiredness are only a few of the early warning signs of CKD.

The CKD is an hospitalized state of the system renal failure, hospitalized individual renal failure, chronic state of the system renal failure, chronic individual renal failure, chronic state of the system renal failure [5].

2 Findings

Overall finding is obvious that different algorithms can appreciate this analysis and review of work being done with, covering, machine and deep learning techniques, and that different algorithms may value this analysis and evaluation of work being done with, covering, machine and deep learning techniques. For this study, we structured our research into three components and two key periods. It uses the R programming language to work with huge datasets. This research paper proposes a method for various machine learning and deep learning algorithms, which has a high level of accuracy. In the future, the proposed study endeavour will be realised in R with a GUI environment, as well as Python.

3 The Dataset's Description

Data Gathering

The UCI machine learning repository compiles the chronic kidney dataset files and then predicts CKD based on the attributes provided. It is necessary to employ both numerical and nominal data. The dataset starts with 103,200 records. Preprocessing, attribute variety approaches, cataloguing, and classification algorithms are just a few of the 33 elements that performance evaluation might use to chronic kidney data.

4 Stacked Autoencoder

An introduction to the autoencoder is required by deep learning architecture based on unsupervised learning for the result of ANN [6]. It is secret input layer encoding and output layer decoding. To based on secret input and output layers of the autoencoder (AE) network are similar to those of artificial neural networks [7]. Each layer of an autoencoder contains a fixed number of neurons. It is an easier to use than feed-forward neural networks. Sparse autoencoders, zero-biased autoencoders, denoising autoencoders, contractive autoencoders, and convolutional autoencoders are all examples of AE [8]. The decoder component is reassemble true message (x'). This fundamental goal is using an AE for extracting key characteristics whilst reducing data size and producing data that are noise-free [9]. The primary idea underlying the aforementioned technique is to look for key patterns in the data [10].

$$C = F\left(w^t x + b\right) \tag{1}$$

$$x' = F\left(wc + b'\right) \tag{2}$$

$$e = \min \sum_{i=1}^{n} \left(x' - x\right)^2 \tag{3}$$

Figure 1 shows that architecture of autoencoder to represented by the input and output layers.

F specifies the artificial neurons been using [11], b the bias value, and w the weights in the input layer and the hidden layers in Eqs. (1) and (2). When using c [12], the variable x' defines the regenerated form of input x. Performance of the model, that also occurs in ANN classifier even though described, is also present in equation, can be reduced via regularisation (4).

$$\min\left[\sum_{i=1}^{n}\left\{\left(x' - x\right)^2\right\} + L1(\text{weight})\right] \tag{4}$$

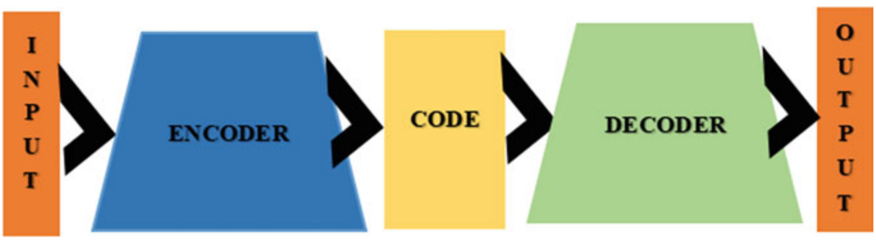


Fig. 1 Autoencoder architecture

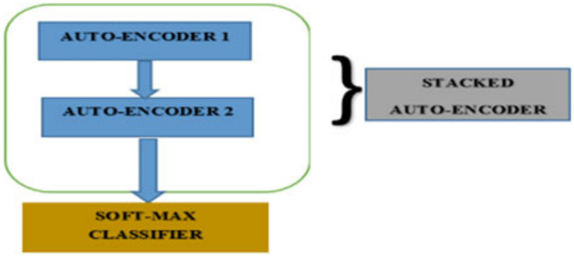


Fig. 2 Process of SAE softmax classifier

L_1 (weight) is the change argument, on whilst signifies regularisation argument in Eq. (4). A hit-or-miss strategy is used to determine the values of these boundaries.

To perform input data categorization, the softmax layer is introduced after the stacked autoencoder. It is a probability-based linear classifier with n inputs that compute the input probability distribution [13]. It simply makes use of the essential qualities, which aids in classification efficiency.

$$F(x_i) = \frac{Exp(xi)}{\sum_{j=0}^{k} Exp(xj)}, \text{ where } i = 0, 1, \ldots k \tag{5}$$

The equation for the softmax function is (5). It calculates the exponential value of the supplied input xi, as well as the sum of all input values' exponential values and the fractional value of the softmax function's output product [6] (Fig. 2).

5 Methodology

The initial goal is to use the chronic kidney disease dataset to make an immediate diagnosis of CKD patients with a high-risk level (Fig. 3).

Phase 1: Deep learnıng algorıthms.

Phase 2: Machıne learnıng algorıthms.

Phase 3: CKD predıctıon performance (Fig. 4).

Phase 1: Deep learnıng algorıthms (Tables 1, 2, and 3).

$$\begin{aligned} \text{Accuracy} &= (20 + 4030)/20 + 0 + 0 + 4030 = 1 = 100 \\ \text{Precision} &= 20/20 + 0 = 1 = 100 \\ \text{Recall} &= 20/20 + 0 = 1 = 100 \\ \text{F1\_ score} &= 2 \times 1 \times 1/1 + 1 = 1 = 100 \end{aligned}$$

$$\text{F1\_ score} = 2 * \frac{(\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{10}$$

Phase 2: Machine learning algorithms (Fig. 5).

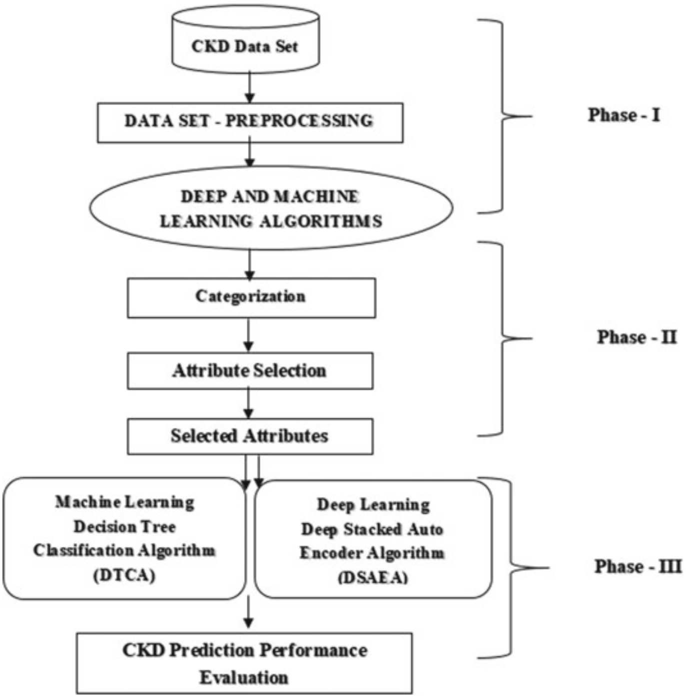


Fig. 3 CKD methodology framework

Decision tree classification algorithm (Table 4).

$$\text{Accuracy} = \left(\frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}}\right) = 100$$

$$\text{Accuracy} = \left(\frac{50 + 4000}{50 + 0 + 0 + 4000}\right) = 100$$

$$\text{TP} + \text{TN} = \text{Correct Prediction and FP} + \text{FN} = \text{Incorrect Prediction}$$

In the table above, TP – true positive and FP – false positive, FN – false negative, and TN – true negative. If a predictive model is established, the accuracy of the predictive model is determined on the basis of accuracy, recall values of the classification matrix, and it is important to verify how accurate it is (Figs. 6 and 7). The fraction of retrieved instances that are significant is precision. It is quantified as

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

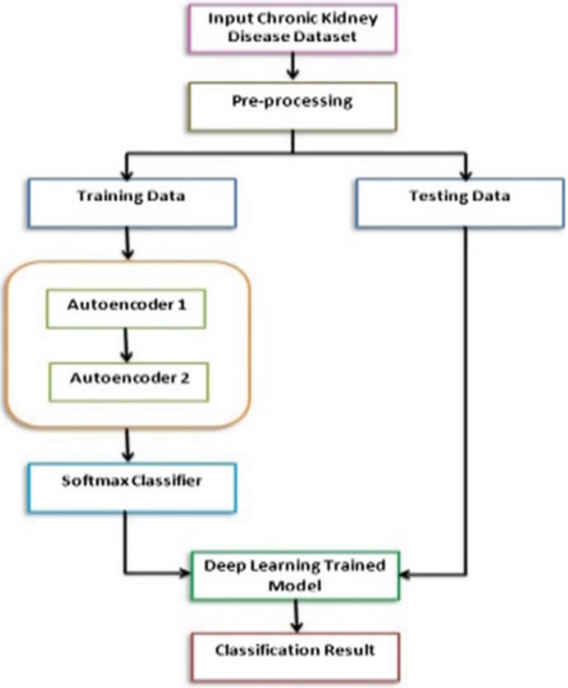


Fig. 4 Methodology framework for classification process of CKD deep stacked autoencoder network

Table 1 Pseudo code for deep stacked autoencoder network

ckdhiddenSize = 10;
ckdautoenc_1 = ckdtrainAutoencoder (X, ckdhiddenSize,…); ckdfeatures_1 = encode (ckdautoenc_1, X);
ckdhiddenSize = 10;
ckdautoenc_2 = ckdtrainAutoencoder (ckdfeatures_1, ckdhiddenSize,…); cdfeatures_2 = encode (ckdautoenc_2, ckdfeatures_1);
ckdsoft_net = ckdtrainSoftmaxLayer (ckdfeatures_2, T, 'Loss_Function', 'cross_entropy'); view (ckdsoft_net)
ckddeep_net = ckdstack (ckdautoenc_1, ckdautoenc_2, ckdsoftnet); ckddeep_net = train (ckddeep_net, X, T);

Table 2 Evaluation metrics of CKD

Evaluation metrics of CKD	Proposed model for DSA (%)
Accuracy of CKD	100
Specificity of CKD	100
Precision of CKD	100
Recall of CKD	100
F1_score of CKD	100

Table 3 Predicted and actual result of CKD

		Actual results	
Predicted results		Positive	Negative
	Positive	200 (TP)	0 (FP)
	Negative	0 (FN)	103,000 (TN)

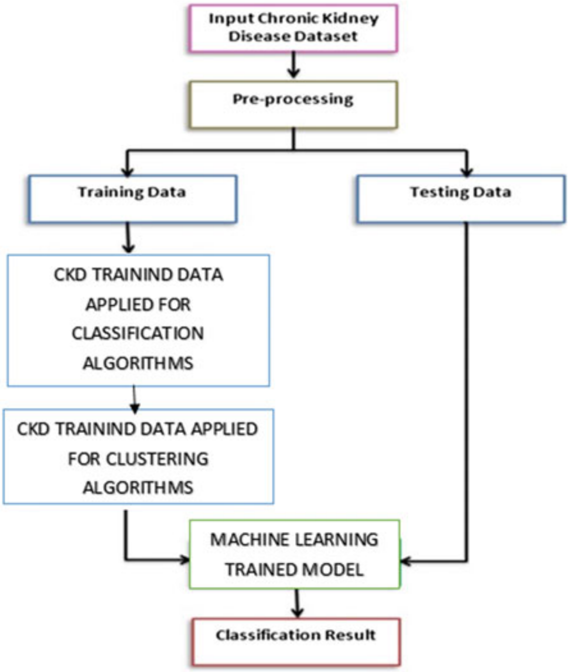


Fig. 5 Methodology framework for classification process of CKD machine learning algorithms

Table 4 Pseudo code for decision tree

Input: Chronic kidney data with selected features
Output: Classified data for decision tree
Step 1: To find a splitting criterion, apply S to D
Step2: Create offspring nodes of *t* if (*t* is not a leaf node)
Step 3: Divide partition *D* into subpartitions for the kids
Step 4: Continue with each segment
Step 5: Come to an end

The recall is a small percentage for total number of associated retrieved instances. It is commonly stated as a percentage. It is measured in terms of (Tables 5 and 6).

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

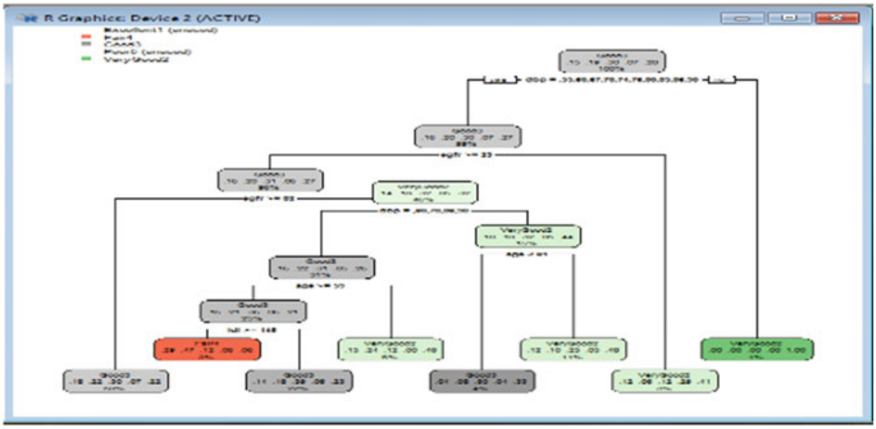


Fig. 6 DTCA using *R* predicted on CKD

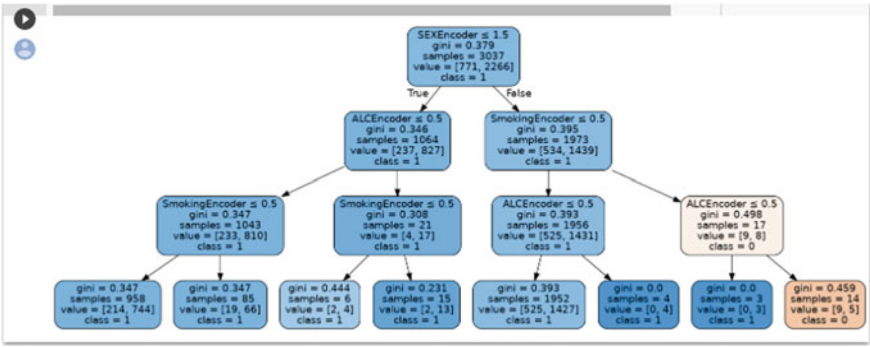


Fig. 7 DTCA using Python predicted on CKD

Table 5 Accuracy performance

CKD performance	Predicted class	
Actual class	TP	FN
	FP	TN
Accuracy (%)		Value

Table 6 Confusion matrix

CKD performance	Predicted class	
	Predicated CKD	Predicated not CKD
Actual class	200	0
	0	103,000
Accuracy (%)		100

The following calculations can be used to evaluate the model's performance (Figs. 8 and 9). The system's ability is used to make accurate forecasts. Table 7 summarises the findings of an accuracy assessment.

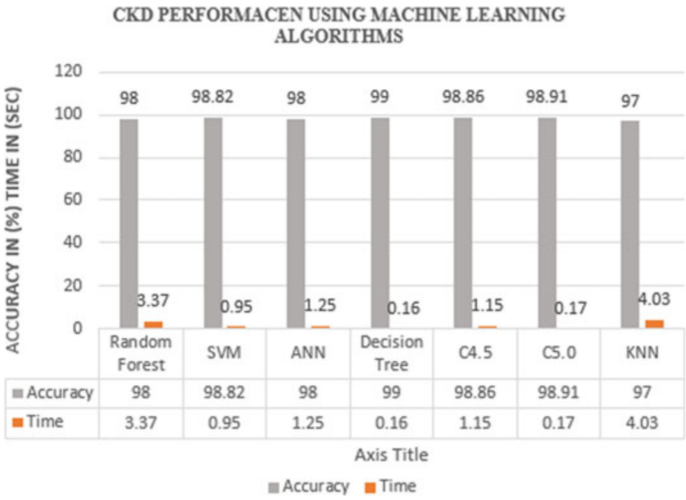


Fig. 8 Accuracy of various machine learning techniques

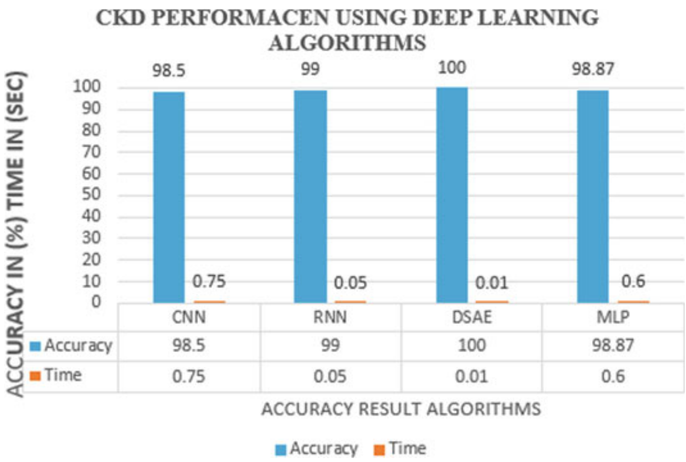


Fig. 9 Performance of several deep learning algorithms in terms of accuracy

Table 7 Accuracy in CKD performance using DCTA

Accuracy of CKD	100%
Sensitivity of CKD	100%
Specificity of CKD	97%

6 Human Role in the IoT-Based Glomerular Chronic Kidney Disease

IoT allows healthcare providers to be more vigilant and proactive in their interactions with patients. Medical equipment such as wheelchairs, defibrillators, nebulizers, oxygen pumps, and other monitoring equipment are tracked in real time using IoT devices with sensors. The purpose of this study was to use an epidemiological concept called stratified hyper filtration, which was stratified by age and sex, to assess the cardiovascular risk associated with GHF in healthy persons.

7 Conclusion

To summarise, the chronic kidney disease plan used to gain access to current machine learning algorithms in healthcare has become an actual clinical diagnosis result, and the use of ml algorithms in collaborative healthcare is generating interest in a promising field for successful outcomes whilst reducing cost. Based on the findings and observations, it can be concluded that the developed deep stacked autoencoder method for such classification of chronic kidney disease would be an excellent process of diagnosing CKD with improved precision, accuracy, specificity, and recall are important in the medical field. In comparison with machine learning, the suggested deep stacked autoencoder and in future, an IoT framework for categorising chronic kidney disease might be a fantastic model to detecting CKD with high accuracy.

References

1. Lakshmanaprabu SK, Mohanty SN, Rani SS, Krishnamoorthy S, Uthayakumar J, Shankar K (2019) Online clinical decision support system using optimal deep neural networks. Appl Soft Comput J 81:105487
2. Ahmed S, Kabir T, Mahmood NT, Rahman RM (2014) Diagnosis of kidney disease using fuzzy expert system
3. Reddy CK, Aggarwal CC (2015) Healthcare data analytics. CRC Press Taylor and Francis Group, ISBN: 13: 978-1-4822-3212-7
4. http://www.webmd.com/urinary-incontinence-oab
5. Bala S, Kumar K (July 2014) A literature review on kidney disease prediction using data mining classification technique. Int J Comput Sci Mobile Comput, IJCSMC 3(7):960–967

6. Khamparia A, Singh A, Anand D, Gupta D, Khanna A, Arun Kumar N, Tan J A novel deep learning-based multi-model ensemble method for the prediction of neuromuscular disorders. Neural Comput Appl 1–13. https://doi.org/10.1007/s00521-018-3896-0
7. Liu W, Wang Z, Liu X, Zeng N, Liu Y, Alsaadi FE (2017) A survey of deep neural network architectures and their applications. Neurocomputing 234:11–26
8. Kannadasan K, Edla DR, Kuppili V (2018) Type 2 diabetes data classification using stacked autoencoders in deep neural networks, pp 2–7
9. Pujari RM, Hajare MVD (2014) Analysis of ultrasound ımages for identification of chronic kidney. First ınternational conference networks and soft computing 380–383
10. Adam T, Hashim U (2012) Designing an artificial neural network model for the prediction of kidney problems symptom through the patient's metal behavior for pre-clinical medical diagnostic, pp. 27–28
11. Bhanodia P, Pandey B, Pandey D, Khamparia A (2019) A comprehensive survey of link prediction in social networks: techniques, parameters and challenges. Expert Syst Appl 124:164–118
12. Chetty N, Vaisla KS, Sudarsan SD (2015) Role of attributes selection in classification of chronic kidney disease patients. IEEE
13. Rosso Retal (2010) Chronious: an open, ubiquitous and adaptive chronic disease management platform for COPD, CKD and renal insufficiency. In: 2010, 32nd annual ınternational conference IEEE EMBS, pp 6850–6853

Evaluating Various Classifiers for Iraqi Dialectic Sentiment Analysis

Nibras Talib Mohammed, Enas Ali Mohammed, and Hafedh Hameed Hussein

Abstract Nowadays, social media outlets involve peoples' opinions, reactions, and emotions. Sentiment analysis classifies the text from those sites into negative or positive polarity. Over the years, a multitude of researchers studied Arabic sentiment analysis but most of them focused on standard Arabic language. However, the Arabic dialects should have much concentration by researchers. Therefore, the main focus of this research is the sentiment analysis of the Iraqi Arabic dialect. The data sourced from Facebook platform (Posts and Comments), the most popular social media site in Iraq. Then, the data passed through several preprocessing steps and weighting methods. The processed data then passed into comparative experiments with six machine learning algorithms including Naïve Bays, Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), Random Forest, and K-Nearest Neighbor (KNN). The results indicated the highest accuracy achieved by Naïve Bays with 81.2%, followed by SVM and LR with 74%, while DT and Random Forest achieved accuracy 64% and 63%, respectively. The worst result was achieved by KNN algorithm of 57%.

Keywords Sentiment analysis · Arabic language · Iraqi dialects · Machine learning

N. T. Mohammed (✉) · E. A. Mohammed
University of Kerbela, Karbala, Iraq
e-mail: nbrass.t@uokerbala.edu.iq

E. A. Mohammed
e-mail: enas.ali@uokerbala.edu.iq

H. H. Hussein
Karbala Education Directorate, Karbala, Iraq
e-mail: hafud_2008@yahoo.com

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_6

1 Introduction

Today, sentiment analysis can be regarded as the trending technology. It offers great enforcement for companies, restaurants, hotels, etc. This technology has a huge impact when it comes to manufacturers, which might require an inquiry to its customers about products [1]. In addition, regarding individuals, sentiment analysis could save time and effort in asking about a product, a movie, etc.

Owing to the huge amount of opinion and emotions existing in social media, analyzing those sentiments has become a significant task in this area to determine the polarity of the written comments [2].

Even though the dramatic improvement of sentiment analysis in the Arabic language, further efforts should be put toward the dialects [3]. Most previous researches focused on standard Arabic [4]. However, Arabic dialects became not only for spoken communications but also for social media text [5]. Dealing with social media written comments is a complex mission due to the informal style and diverse spelling forms. Consequently, extra preprocessing is needed.

In this research, a model of Iraqi dialect sentiment analysis will be built based on machine learning algorithms. Several machine learning algorithms will be exploited to predict the polarity of the given text. The utilized algorithms are Naïve Bays, Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), Random Forest, and K-Nearest Neighbor (KNN) to compare the performances.

The rest of the paper was organized as followed: The related works is presented in Sect. 2. The methodology of this work is illustrated in Sect. 3. In Sect. 4, the results of this work were listed and discussed. In Sect. 5, the paper was concluded with the key findings.

2 Related Works

A multitude of studies has been conducted in the field of sentiment analysis. For instance, the authors in [6] compared between Bayesian Rough Decision Tree and Decision Tree algorithms. In this work, Facebook and movie review datasets were exploited. The results revealed that Bayesian Rough Decision Tree outperformed the other algorithm for both datasets with threshold values 0.01 and 0.02 for the Facebook dataset, and 0.01 and 0.04 for movie review datasets.

In terms of Arabic sentiment analysis, many approaches have been proposed by researchers in this area. The work of [7] concentrated on building an Arabic dataset from diverse sources such as the Twitter platform, hotel reviews, and restaurant reviews from different web sites. Then, a comparison has been conducted among Naïve Bayes, Support Vector Machine, and Decision Tree. The best results were gained by Support Vector Machine with a product attraction dataset. The authors of [8] tested different features such as sentiment lexicon, n-grams, and Ara Vec word embedding, which outperformed other features. Furthermore, they examined the performance of three machine learning algorithms including random forest, ridge classification, and support vector classification with L1 and L2 penalties as well as a combination of all of them. The latter algorithm with L1 outperformed the other tested algorithms.

However, Iraqi dialect sentiment analysis suffers from a lack of work done in this area. In [9], the researchers analyzed Facebook comments about the Iraqis government and Iraqis politicians. The proposed model comprises of K-Nearest Neighbor, AdaBoost, and Naïve Bayes ensemble machine learning algorithms. For the latter algorithm, Bernoulli and Multinomial models were applied. The results of this research depicted that the highest accuracy is given by Naïve Bayes with a Multinomial model. Further Iraqi dialect research is [10] who built the (IAEDS) Iraqi emotion recognition dataset. Then, the dataset was tested by five algorithms including J48, Naive Bayes, SMO, ZeroR, and Multinomial Naive Bayes for text. The last two algorithms presented the worst results. Moreover, another classifier was proposed in this work called prediction by partial matching PPM, which outperforms the other explored algorithms.

In [11], an experimental study was done to evaluate well-known algorithms (SVM and Random Forest) with other deep learning models (CNN) in classifying the Algerian Dialect texts (opinion and Emotions). The result showed that the CNN model gave better accuracy than other models.

The authors in [12] have introduced a framework to detect sentiment polarity in Arabic tweets datasets (general Arabic and Jordanian dialect texts). The proposed framework was used several machine learning classifiers (SVM, Naive Bayes, Multilayer perception (MLP), and logistic regression-based). The classifiers showed better performance by using a combination of n-gram features.

In [13], a semi-supervised approach was implemented for sentiment polarity classification of Jordanian dialectal reviews. Well-known classifiers (SVM, NB, Random Forest, and KNN) were used and combined with dialectal lexicons to determine the best classifiers. Three evaluation methods correlation-based feature selection, Principal components analysis, and SVM features) were used in this paper to select the features with which the classifiers can perform the best. The SVM obtained the highest accuracy with 92.3%.

However, this research focuses on comparing a set of six classification algorithms and proving the accuracy of the best ones for the Iraqi dialect sentiment analysis.

3 Methodology

Due to the obstacle of limited Iraqi datasets, the first step of this research was collecting the Iraqi dialect dataset. Cleaning data from unwanted characters, stop words, and stemming data was a further phase of this work. Selecting the significant features by the use of the Term frequency-inverse document frequency (TF-IDF) weighting scheme. Training the classification model was the latter step before testing and validating the model phase. The phases of this work are illustrated in Fig. 1.

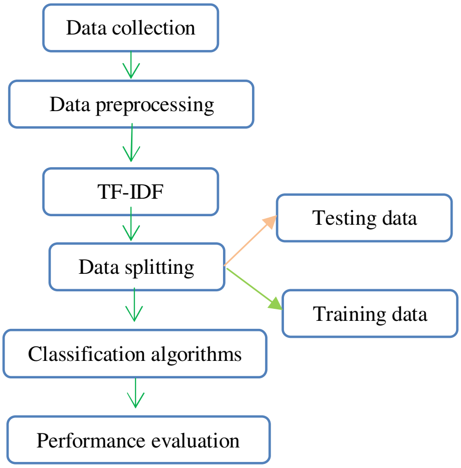


Fig. 1 Iraqi sentiment analysis model

3.1 Dataset

Iraqi people speak several languages but the Arabic language is the formal language. Arabic dialects are closer to standard Arabic with some differences in writing and it does not have a dictionary and rules. In social media, most users write posts/comments in a dialect where users have the freedom to express their opinion with special writing methods since the writing and spelling differing from one person to another. Therefore, processing the Arabic dialects is more difficult than the standard Arabic. In this research, the Iraqi dialect dataset has been collected from Facebook, the most popular social networking platform in Iraq. The dataset extracted from both Facebook posts and comments from eight famous Iraqi news channels pages. The Facepager platform was exploited for extracting those data. 1080 comments and posts have been collected; then, Tagging of the dataset was performed manually to their sentiment (positive or negative). The data divided into 540 comments and posts for each category. A sample of the Iraqi Arabic dialect dataset is illustrated in the Table 1.

Table 1 Examples of sentences in the dataset and their mood labels

Arabic sentence	English sentence	Sentiment label
Fe haditha ghariba dub muftaris unqith tiflan fe al thalitha min al omr	In a strange incident … # A ferocious bear rescues a three-year-old child!	Negative
Hatha al rajul shareef watani khadam watanahu al Iraq beikhlas wa dafaa anhu fi kull al horob wa kan mokhlis wa lays khaen mthl al khawana	This man is honest and patriot, he served his homeland, Iraq wholeheartedly, and defended him in all wars. He was loyal to his country and not a traitor to his country like the rest of the traitors	Positive

Fig. 2 Preprocessing steps

3.2 *Preprocessing*

Owing to the noise and unstructured nature of the data coming from the Internet, the data should be processed by the following steps depicted in the Fig. 2.

Natural language processing (NLP) covers human language understanding and manipulation. Thus, NLP processes will be applied to our dataset to clean it from noise.

Firstly, the tokenization process has been applied, which comprises splitting each comment into tokens (single words). Cleaning comments from unwanted characters such as special characters, digits, English words. Stop words are not affecting the sentiment decision [14]. Besides, stop words affect the accuracy and time of processing. Thus, a further preprocessing step was removing Arabic stop words from the dataset. The last preprocessing step is stemming, which meant grouping some related word in terms of meaning or structure into one root as heavy stemming, and remove prefix, suffix, and infix as light stemming. In this paper, Information Science Research Institute (ISRI) Arabic stemming has been carried out to extract roots of the words as light stemming. This stemmer avoids invalid or insensible roots production due to the context-sensitive nature of this stemmer [15].

3.3 *Feature Selection*

Feature selection in machine learning can be regarded as a crucial process since maximum accuracy can be achieved by selecting the best set of features.

The words that have a significant impact to differentiate between dissimilar categories represent the set of features in text classification [16]. In this work, Term Frequency-Inverse Document Frequency (TF-IDF) has been exploited for this purpose since this scheme regards the significance of the term or word in the whole corpus instead of considering it in only a document. The value of TF can be calculated by the term (t) with the frequency ft in a given document (d). IDF calculation can be performed by the frequency Fd of document d, which contains the term (t) divided by the number (N) of the document in the whole dataset as depicted in (Eq. 1). The value of TF-IDF, which considered weight for words in corpus calculation, can be obtained by multiplication of the value of TF and the value of IDF as shown in (Eq. 2) [17].

$$\mathrm{IDF}(t) = \log\left(\frac{f_d}{N}\right) \tag{1}$$

$$\text{TF-IDF}(t, d) = \mathrm{TF}(t, d) \times \mathrm{IDF}(t) \tag{2}$$

3.4 Training and Testing

To train and test the preprocessed labeled dataset, the data is spliced into 80% and 20% for training and testing, respectively. Then, several classification algorithms were exploited to achieve the best accuracy. The elected classification algorithms were based on the classifiers, which were frequently utilized in the research area of sentiment detection [1, 18–21]. The classifiers comprise Naïve Bays, Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), Random Forest, K-Nearest Neighbor (KNN) to compare their performance in categorizing the comments into positive or negative.

3.5 Evaluation Metrics

The metrics to evaluate the performance of utilized classifiers included accuracy and F-measure. Those metrics can be calculated by the use of true-positive (TP), false-positive (FP), true-negative (TN), and false-negative (FN) metrics.

TP states to the positive instance, which is classified as positive. FP indicates the negative instances, which are predicted as positive, TN refers to the negative instances and predicted as negative, and FN indicated the rate of the instances predicted as negative incorrectly.

Accuracy represented the rate of the correctly predicted instances.

The Eqs. (3, 4, 5, 6) below are criteria of evaluations of the classification algorithms.

$$\mathrm{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{3}$$

$$\mathrm{Precision} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP}) \tag{4}$$

$$\mathrm{Recall} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN}) \tag{5}$$

Table 2 Performance results of the classifiers

Classifiers	Accuracy (%)	F-Measure (%)	Precision (%)	Recall (%)
Naïve Bays	81	82	84	81
SVM	74	78	70	89
Logistic regression	74	77	74.8	79
Decision tree	64	49	97	33
Random forest	63	74	98	33
KNN	57	35	90	21

$$\text{F-Measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{6}$$

4 Results and Discussion

The performance results of the utilized classifiers are listed in Table 2 using the metrics mentioned in the previous section. As illustrated in the table that Naïve Bays achieved the highest results in terms of all performance metrics, followed by SVM and LR with the same accuracy with 74% and converging values of the other metrics. DT and Random Forest results were converging also in terms of all performance metrics where accuracy values were 64% and 63%, respectively. However, the worst results were achieved by KNN algorithm with accuracy 57%.

5 Conclusion

The use of data from Microblogging sites such as Facebook and Twitter has become popular in the sentiment analysis area. Thus, our work depends on the dataset collected from the Facebook post. Then, the data was preprocessed and weighted to be ready to enter the classification algorithms. Popular classification algorithms have been applied including Naïve Bays, SVM, LR, DT, Random Forest, KNN, and the findings depicted that Naïve Base outperformed other classification algorithms with accuracy 81% and F-measure 82%, while the least accuracy accomplished by KNN.

References

1. Altowayan AA, Tao L (2016) Word embeddings for Arabic sentiment analysis. Proceedings—2016 IEEE international conference big data, big data 2016, pp 3820–3825. https://doi.org/10.1109/BigData.2016.7841054
2. Sabbah T et al (2017) Modified frequency-based term weighting schemes for text classification. Appl Soft Comput J 58:193–206
3. Al-Kabi MN, Rabab'Ah AM, Al-Ayyoub M, Jararweh J. Evaluating sentistrength for Arabic sentiment analysis. In: 2016 7th international conference on computer science and information technology (CSIT) (IEEE)
4. Duwairi RM, Ahmed NA, Al-Rifai SY (2015) Detecting sentiment embedded in Arabic social media—a lexicon-based approach. J Intell Fuzzy Syst 29:107–117
5. Al-Azani S, El-Alfy ESM (2017) Using word embedding and ensemble learning for highly imbalanced data sentiment analysis in short Arabic text. Procedia Comput Sci 109:359–366
6. Alatabi HA, Abbas AR (2020) Sentiment analysis in social media using machine learning techniques. Iraqi J Sci 61:193–201
7. Algburi MA, Mustapha A, Mostafa SA, Saringatb MZ (2020) Comparative analysis for Arabic sentiment classification. In: Communications in computer and information science, vol 1174 CCIS. Springer, Cham, pp 271–285
8. Badaro G, et al (2018) EMA at SemEval-2018 task 1: emotion mining for Arabic, pp 236–244. https://doi.org/10.18653/v1/s18-1036
9. Habeeb LA (2018) Sentiment analysis for iraqis dialect in social media. Iraqi J Inf Commun Technol 1:24–32
10. Almahdawi AJ, Teahan WJ (2019) A new Arabic dataset for emotion recognition. Adv Intell Syst Comput 998:200–216
11. Moudjari L, Akli-Astouati K (2020) An experimental study on sentiment classification of Algerian dialect texts. Procedia Comput Sci 176:1151–1159
12. Hawalah A (2019) A framework for Arabic sentiment analysis using machine learning Ahmad Hawalah To cite this version: HAL Id: hal-02300717 a framework for Arabic sentiment analysis using machine learning classifiers
13. Al-Harbi O (2019) Classifying sentiment of dialectal Arabic reviews: a semi-supervised approach. Int Arab J Inf Technol 16:995–1002
14. Goebel R (2018) Data mining
15. Mulki H, Haddad H, Gridach M, Babaoğlu I (2018) Tw-StAR at SemEval-2017 task 4: sentiment classification of Arabic tweets, pp 664–669. https://doi.org/10.18653/v1/s17-2110
16. Chen J, Yan S, Wong KC (2020) Verbal aggression detection on Twitter comments: convolutional neural network for short-text sentiment analysis. Neural Comput Appl 32:10809–10818
17. Alessa A (2018) F. M. Tweet classification using sentiment analysis features and TF-IDF weighting for improved flu trend detection. In: Perner P (eds) Machine learning and data mining in pattern recognition. MLDM 2018. Lecture notes in computer science. Springer, Cham, pp 174–186
18. Abdulla NA, Ahmed NA, Shehab MA, Al-Ayyoub M (2013) Arabic sentiment analysis. Jordan conference applied electrical engineering and computing technology, vol 6, pp 1–6
19. Aly M, Atiya A (2013) LABR: a large scale Arabic book reviews dataset. ACL 2013—51st annual meeting association for computational linguistics proceedings conference, vol 2, pp 494–498
20. Chandrashekhar M, Ibrahim RW, Obaid AJ, Meshram SG, Meshram A, Abd El-Latif AM (2020) Fractional chaotic maps based short signature scheme under human-centered IoT environments. J Advan Res. ISSN 2090-1232. https://doi.org/10.1016/j.jare.2020.08.015
21. Obaid AJ, Chatterjee T, Bhattacharya A (2020) Semantic web and web page clustering algorithms: a landscape view. EAI Endorsed Trans Energy Web 8(33)

Sentiment Analysis of Software Project Code Commits

Archana Patnaik and Neelamadhab Padhy

Abstract Sentiment analysis is used to analyze the impact of Git commit on various open-source Java projects. In order to understand the positive, negative, and neutral impact of sentiment on developer community, we have reviewed many research paper based on real-time projects. It is observed that emotion is having a wide impact on software product quality by considering different factors like software requirement, team management, and work distribution which contributes toward negative commit comments. Code commit analysis is either done on daily or weekly basis by considering 8 classifications of emotions. In order to do a detailed survey, different open-source projects like Eclipse, JEdit, ArgoUML, and JUnit are considered for commit analysis based on software refactoring activities. In order to give an exact statistical result on sentiment analysis, different machine learning and deep Learning classifiers can be used. Based on the survey, it is also concluded that code refactoring is highly influenced by positive or negative impact of developer's emotions.

Keywords Code commit · Sentiment analysis · Emotions · Software quality · Data repositories · Software refactoring

1 Introduction

Sentiment analysis is the study of expression that is used to identify human emotions, attitude, choice, and attitude toward a particular real-time activity. It is also defined as the automated mechanism of understanding of feelings based on contextual and textual analysis. It can be classified into three types that is negative, positive, and neutral sentiments. It emphasizes on various software engineering tasks using various manual and automated tools. In order to improve the internal and external quality of

A. Patnaik (✉)
School of Engineering and Technology Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India
e-mail: archanapatnaik@giet.edu

N. Padhy
Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India
e-mail: dr.neelamadhabpadhy@giet.edu

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_7

the software product, it is very essential to study the behavior and mind set of the developer during software development life cycle. It is used in investigating different activities like selection of programming language, implementation of software code, code execution, introduction of bug, finding error or complexities, and fixing of bug. In order to do the details analysis, many researchers have considered emotion as the major parameters. Emotion analysis is the process of identifying and analyzing the expression of a person based on the contextual data. Developer's emotion affects the productivity and task quality. In order to give the in depth analysis, we have to come across 8 classification of emotion which includes sadness, disgust, anticipation, trust, anger, fear, and joy. It may have a positive or negative impact on team division, task quality, and job satisfaction of a software engineering. The consideration is done by considering many real-time Java projects. Emotion score [1] is calculated either on daily or monthly basis. It is detected by using facial expression, body movement, and working strategy of an individual. In our work, we are going to study the effect of emotions on different development activities. Code commit is illustrated as the specific changes made to the dataset of project which collected from private GitHub repositories. It occurs automatically or manually between the initial and final stage of project execution. In order to execute it practically, we have to consider the training and prediction elements of a hybrid model. It is exhibited in the structural level in order to control the entire work flow of the system.

Code refactoring is also directly related to emotional mind study of the programmer during the code execution. Based on the previous work, it is observed that refactoring activities like extract method, extract interface, move attributes, move class, move method, pull up method, rename class, and inline method are analyzed using the SentiStrength tool. The sentiment percentage is calculated by considering the probable value of different code based activities. Most of the time negative sentiment becomes the result of the software code commit analysis. If the probable value is less than negative value, then in that case sentiment percentage increases from the average value. By using GitHub pull request flow, bug identification and fixing are done at developer level. By doing the statistical correlation analysis on number of developer and size of project in term of LOC, it is concluded that 14% of developers are having negative sentiments if the line of code is increased on regular basis. It also has a wide impact on cost efficiency, time, space, and code complexity by effecting the software quality which is surveyed by quality assurance and assessment. The end point parameters are code size, polarity which may be positive, negative, or neutral and collection of customer feedback. Sometimes mood variation also affects the result analysis part of the final product.

In this paper, we have done a detailed survey of sentimental analysis on software code commit. We have also come across 8 classifications of emotions. The researcher has considered different real-time dataset and determined the emotion score based on which sentiment of the developers is analyzed. Some researchers have also used tools like SentiStrength and SentiCR for performing various activities analysis. It is observed that mood variation also varies based on number of lines of code calculated in terms of LOC and KLOC. Based on our literature survey, we have framed three different research questions:

RQ 1: What are the different levels of emotions that effect developing activities during software development life cycle?

Based on different experimental and statistical analysis, we have observed that there are three levels of emotion that it positive, negative, and neutral. The emotional score defines the levels of sentiments and always it tends toward negative direction.

RQ 2: Does code refactoring activities have an impact on developer's sentiments?

Refactoring occurs due to change in internal structure of code without modifying its external behavior. Sentiment of the developer is also affected by the alteration activities. Emotion analysis is done by considering code commit comments.

RQ 3: Does variation in human emotion effects the software product quality?

The effects of negative emotion widely contribute a lot in reducing software quality which also decreases its maintainability and reliability. During emotion score analysis, we will test all the categories like happiness, sadness, anger, love, fear, etc.

In Sect. 1, we have discussed the introduction of sentiment analysis which is followed by literature survey in Sect. 2. Based on the related work, we have also framed different research questions. Section 3 provides the detailed description about the research background. In Sect. 4, we have proposed and organized research framework. Section 5 provides the effects of emotion on software products which is shown in the form of result analysis. In Sect. 6, we have concluded our work and its future scopes are stated.

2 Related Works

Sentiment analysis [2] extracts different attributes that is collected from Git commit comments of different open-source Java projects. In order to determine the power of sentiments on developer's emotion, different machine learning algorithms are used. Consequently, it highly influences the software product by assuring its quality. Researchers investigate the emotions by performing the textual and contextual analysis of code commits. Analysis of different factors like use of relevant coding, selecting programming language, submitting daily, and weekly progress activities contributes a lot in improving software quality. The work done in this paper [3] illustrates the neutral, negative, and positive comments based on software code commits. In this work [4], analysis of developer is done by considering 998 data of GHTorrent dataset which is collected from GitHub sites. In order to understand eight different classifications of emotions, various commit messages are taken as evaluation data. Based on the result, analysis 45% of commit messages displays trust as the most favorable type of emotion find during weekly data analysis. Development activities are affected by human emotions variation which varies in between bug fixing and bug introducing. Quantitative studies [5] are done on three open-source projects at class and method level. It is observed that 24,000 commit changes are made during the lifecycle of software project. Based on the expressed emotion, bug is identified and corrected by using commit methods. The result shows the positive or negative impact of bug on sentiment of the software developer. The research goal [1] is to identify the variation of emotions in male and female during various software engineering tasks. They have used sentiment tools for understanding neutral, negative, and positive emotions. In order to resolve the gender bias, issues different code segments are reviewed by considering lines of code. Based on comparative studies, it is concluded that females are less expressive than males as the negative work atmosphere effects the emotions of the developer. It is also observed that females are more neutral than males. Recent studies of the author [6] focus on different aspects of software engineering by correlating the emotions and sentiments of the software team members. The exact analysis can be done by using different human evaluator tools like NLTK and SENTISTRENGTH which give a trained project review by conducting various comparative studies. They have repeated the testing on seven open-source projects by using stack overflow and issue tracker tools. This paper [7] illustrates the effects of code refactoring on emotions which working on real-time projects. Analysis is done on 60 Java projects where 15 refactoring activities are performed. Out of 4891 instances, 3171 refactoring commits are done by working on different types of data. The result shows that five refactoring activities give proper outcome-based positive results. 10% of software developer contributes toward the negative sentiment of the projects. The software developer [7] considers the change in commit log during the bug analysis while performing software testing. The commits are classified under three valid categories like medium, small, and large. Weekly sentiment analysis is done on 28,466 projects from GitHub data. The statistical evaluation shows the correlation between number of file s modified and sentiments of developers are directly proportional to each other. Their finding shows that negative sentiments is 10% more than positive, and Tuesday gives maximum negative commit out of remaining days of the week. In this work [8], the customized sentiment tools SentiCR tool is used for software product review interactions on project datasets. The poor performance of the team members shows the negative impact of emotions on developmental activities. 2000 review comments are surveyed on training and testing data for the hybrid model. The performance evaluation is done by using gradient boosting classifiers which gives highest recall and precision values of 58.4% and 67.8% which contributes toward negative review comments. Software development [9] process depends on connections productivity on 197 KLOC analyzing the mood variation on each developer. In software projects, analysis 34 KLOC gives 10.47% of negative commits, and 24.02% contributes toward positive commits. GitHub pull request shows the mood variations within the range of 78–268 K understanding the feedback developer sentiments. The behavioral changes may not occur due to sentiment but the reason of mood variation is negative comments. In the work stated by the authors [10], community smell is predicted by based on the developer's emotions. Their work focuses on construction of developer-oriented smell removal community, where different refactoring methods are used. Based on the survey [11], it is observed that sentiments analysis related to the software bugs is also affecting the GitHub commit. In order to analyze the detailed statistics, difference between bug and correct comments is analyzed by the researcher.

3 Research Background:

The overall research background is segregated into three major categories which includes

A. Sample Source Projects

In order to choose sample projects, we have searched different software engineering repositories. We have also referred to GitHub from where many industrial experts have stored the data based on real-time project analysis. Many types of projects like Junit, ArgoUML, Eclipse, Open stack, LibreOffice, Android, and other Java-based project are collected by different researcher for persuading efficient research. The developers have used tools like Boa mining tools and many other data extractor which is used to collect data from the above projects. Out of 150 GitHub projects, we have selected four randomly by considering its probable value which is useful to calculate the emotional score for sentiment analysis. After collecting the raw data, it further undergoes data preprocessing and removes biased and null values. By using various data extractor algorithms of machine learning, we can select the relevant features. The processed data enter into the next phase for score analysis. There some technical tool used to NLTK, SENTISTRENGTH which is used to collect the emotion score from the dataset.

B. Refactoring Activities

Code refactoring is the process of modifying the internal structure of code without affecting the external output. It is done to reduce the code smell which occurs due to length code, code complexity which occupies more space and more time. Different types of smells are bloater, large class, long method, god class, feature envy, duplicate code, etc. In order to compensate the code smell, researchers have used methods like extract class, extract method, push method, inline method, and inline class. For efficient analysis of developer's emotions, some refactoring activities are organized move method, move class, move attributes, push down method, and push down attributes. In order to exhibit project cloning and commit level extraction, RefTypeExtractor and RefactoringMiner are used. The false positive rate is evaluated from the real-time data collected from the projects, where individual score is compiled to result the overall score. The output analysis is done to differentiate between commit message and refactoring commit message based on clone GitHub projects. The internal modification of code may have a positive or negative impact on software team member emotions.

C. Sentiment Analysis

Sentiment analysis is the quantitative mood value analysis of a developer during the process of software development lifecycle. It surveys the code snippet and provides a proper review in three different categories that are positive, negative, and neutral. It determines the polarity by using different tools like Alchemy, SENTISTRENGTH, and NLTK. The concept of reusability also widely affects the mood score resulting in variation of emotion of a person. The evaluation can be either done manually or automatically by using different automated tools. Based on the estimation of researchers [3], neutral scores vary from −1 to 1, positive scores vary from 1 to 5, and negative scores range lies between −5 and −1. In paper [7], polarity and strength of the sentiment are detected by using NLTK tool which is widely used in identifying the software artifacts. It also plays a vital role in determining the time duration of project completion. The segregation of commit comment is done based on time slot assigned to each individual. It is observed that probable value 0.066 gives most accurate result as compared to other values.

4 Proposed Research Framework

In the Fig. 1, we have proposed a hybrid model and framed three research questions. First phase explains about collection dataset from repositories like GitHub. Data are fetched from many real-time projects like Eclipse, ArgoUML, LibreOffice, Open stack, etc. In order to analyze the sentiment of software code, commit message emotional score is calculated from the projects. Data extractor algorithm of machine learning includes various feature selection and feature extraction methods. It is used to remove the bias, error, and null value present in the projects and generate a finely refined dataset with relevant data. In third phase, the final dataset is collected and is

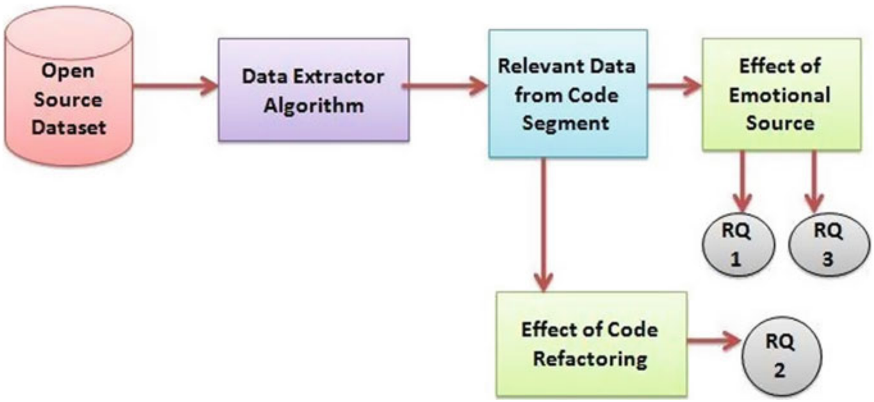


Fig. 1 Research model

saved for code commit comment analysis. Some part of the entire project is segregated into small parts and tested for emotion analysis which is done by considering the mood value. In phase four, we mainly focus on emotional score which is collected from the source probable value. The mood variation of an individual can be easily evaluated either from the mood value or the emotional score which is framed in research question 1 and 3. There are different score of positive [1–5], negative [−5, −1], and neutral [−1, 1] emotions. Code review is again divided into two categories that commit message and refactored commit message. In phase six, we are going to identify the effect of refactoring activities on emotional activities based on which research question two is framed. It focuses on changing the internal code structure without affecting the results which is also having a wide impact on developer's emotions. The detailed description about refactoring activities is illustrated in second research question which shows emotional variation and its relationship which is classified by the time slot. The previous work identifies total commit messages [7] for 60 projects which is figure out as 615,625.

5 Result Discussions

Our result focuses on different aspects of sentiment, and emotion analysis is explained with the help of three research questions. It also illustrates about various manual and automated tools used for calculating the emotional score of a developers.

RQ 1: What are the different levels of emotions that effect developing activities during software development life cycle?

Emotion is defined the evaluation of positive and negative sentiment which is effected at commit level. Commit occurs when there is sudden change in developer's emotions, software activities are also influenced. In order to understand different level of emotions, we have to consider many real-time projects in which probable value lies between the ranges of 0 to 10. Based on the p- value, the emotional score of a developer lies within the ranges of 9.5 to 10.5. It is observed that the p-value remains constant after 6 which show no variation in the emotions. In this way, we can understand different levels of emotions and its changes that occurs during the software development activities (Fig. 2).

RQ 2: Does code refactoring activities have an impact on developer's sentiments?

Based on the literature survey, we have observed that code refactoring can either have a positive or negative influence on developer's emotions. With the help of RefType-Extractor tools, we can clone the inputs given to the projects and can extract total numbers of commit message present at Git commit levels. The effort given to modify the internal code can be either manual or automatic. The output of SentiStrength tools indicates different stages like code development, implementation, writing comments, understanding code, refactoring, etc. (Table 1).

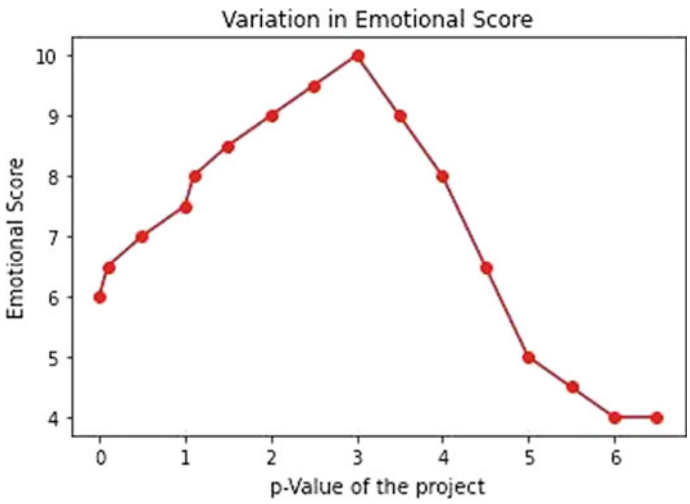


Fig. 2 Variation in human emotions

Table 1 Refactoring activities of software projects

Sl. No.	Refactoring activities	Projects
1	Extract method	JEdit
2	Inline method	Eclipse
3	Extract class	ArgoUML
4	Move method	LibreOffice
5	Move class	Open stack

RQ 3: Does variation in human emotion effects the software product quality?

Human emotion is classified into different categories like joy, anger, sadness, love, hatred, anticipation, happiness, etc. Generally is analyzes the gap between bug introduction and bug fixing. Maximum positive and negative sentiments vary on weekly basis, and the percentage of commit [12] is very high on Tuesday and Wednesday and is very low on Saturday. The variation in emotion widely effects the software quality as the negative sentiments decreases the product quality by improper code design (Fig. 3).

6 Conclusion

In this work, we have surveyed different aspects of sentiment analysis using different open-source real-time projects. In order to understand the positive and negative impacts, we have consider emotional score as the parameters. Different levels of variation can be determined by the variation of emotional scores which lies between

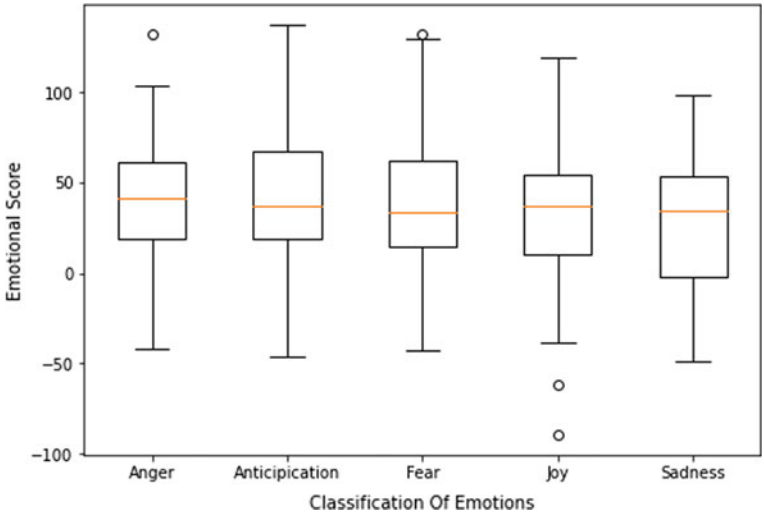


Fig. 3 Classification of emotion

9.5 and 10.5 points. Based on the survey, it observed that emotion is directly affecting the level of refactoring activities which is identified by emotional scores. The graphical analysis shows the relationship between number of commit and emotional scores before and after refactoring which widely affect the quality of the mobile application software. In order to increase the efficiency of the project classification of various types of emotions like anger, anticipation, fear, joy, and sadness are considered. The detailed parametric evaluation by considering the source code data will be reflected in our future work.

References

1. Paul R, Bosu A, Sultana KZ (Feb 2019) Expressions of sentiments during code reviews: male versus female. In: 2019 IEEE 26th international conference on software analysis, evolution and reengineering (SANER). IEEE, pp 26–37
2. Ronchieri E, Juric R, Canaparo M (2019) Sentiment analysis for software code assessment. In: 2019 IEEE nuclear science symposium and medical imaging conference (NSS/MIC). IEEE pp 1–2
3. Guzman E, Azócar D, Li Y (May 2014) Sentiment analysis of commit comments in GitHub: an empirical study. In: Proceedings of the 11th working conference on mining software repositories, pp 352–355
4. Venigalla ASM, Chimalakonda S (May 2021) Understanding emotions of developer community towards software documentation. In: 2021 IEEE/ACM 43rd international conference on software engineering: software engineering in society (ICSE-SEIS). IEEE, pp 87–91
5. Islam MR, Zibran MF (2018) Sentiment analysis of software bug related commit messages. Network 740:740

6. Jongeling R, Sarkar P, Datta S, Serebrenik A (2017) On negative results when using sentiment analysis tools for software engineering research. Empir Softw Eng 22(5):2543–2584
7. Singh N, Singh P (Dec 2017) How do code refactoring activities impact software developers' sentiments? An empirical investigation into GitHub commits. In: 2017 24th Asia-Pacific software engineering conference (APSEC). IEEE, pp 648–653
8. Ahmed T, Bosu A, Iqbal A, Rahimi S (Oct 2017) SentiCR: a customized sentiment analysis tool for code review interactions. In: 2017 32nd IEEE/ACM international conference on automated software engineering (ASE). IEEE, pp 106–111
9. Freira M, Caetano J, Oliveira J, Marques-Neto H, Analyzing the impact of feedback in GitHub on the software developer's mood
10. Huang Z, Shao Z, Fan G, Gao J, Zhou Z, Yang K, Yang X (2021) Predicting community smells' occurrence on individual developers by sentiments. arXiv preprint arXiv:2103.07090
11. Huq SF, Sadiq AZ, Sakib K (Feb 2020) Is developer sentiment related to software bugs: an exploratory study on github commits. In: 2020 IEEE 27th international conference on software analysis, evolution and reengineering (SANER). IEEE, pp 527–531
12. Sinha V, Lazar A, Sharif B (May 2016) Analyzing developer sentiment in commit logs. In: Proceedings of the 13th international conference on mining software repositories, pp 520–523
13. Hajhmida MB, Oueslati O (2021) Predicting mobile application breakout using sentiment analysis of facebook posts. J Inf Sci 47(4):502–516
14. Sagar PS, AlOmar EA, Mkaouer MW, Ouni A, Newman CD (2021) Comparing commit messages and source code metrics for the prediction refactoring activities. Algorithms 14(10):289
15. AlOmar E, Mkaouer MW, Ouni A (May 2019) Can refactoring be self-affirmed? An exploratory study on how developers document their refactoring activities in commit messages. In 2019 IEEE/ACM 3rd international workshop on refactoring (IWoR). IEEE, pp 51–58
16. Alomar EA, Peruma A, Mkaouer MW, Newman CD, Ouni A (2021) Behind the scenes: on the relationship between developer experience and refactoring. J Softw: Evol Process e2395
17. Patnaik A, Panigrahi R, Padhy N (March 2020) Prediction of accuracy on open source java projects using class level refactoring. In: 2020 international conference on computer science, engineering and applications (ICCSEA). IEEE, pp 1–6
18. Patnaik A, Padhy N (2021) A hybrid approach to identify code smell using machine learning algorithms. Int J Open Sour Softw Processes (IJOSSP) 12(2):21–35

Internet of Robotic Things: Issues and Challenges in the Era of Industry 4.0

Geetika Madaan, H. R. Swapna, Sanjeet Singh, and D. Arpana

Abstract IoT is growing at a fast pace, and billions of devices are now associated with the amount expected to reach trillions in the coming years. The Internet of things and the individual systems participate closely in launching the fourth industrial revolution and form alliances and develop the goods of the next generation. The foundation of Industry 4.0 is the transforming innovations. The convergence of robot and IoT agents contributes to Internet of Robotic Things concept, where creativity in automated devices generates different possibilities, both in business and science. It covers a range of sectors, including agriculture, manufacturing, health, education, and surveillance though the application of various technologies. The study discusses the new Internet of Robotic Things developments, which have an influence on the area of health, science, agriculture, manufacturing, education, and surveillance and the key open problems of introduction of robot technology into intelligent spaces. Internet of Robotic Things technology and frameworks are often addressed to highlight their effect on daily life and to promote more study on remote and automatic applications.

Keywords Cyber-physical systems · Education · Health care · Industry 4.0 · Internet of Things

G. Madaan (✉)
University Centre for Research and Development (UCRD), Chandigarh University, N95, Chandigarh State Hwy, Ludhiana, Punjab 140413, India
e-mail: geetumadaan@gmail.com

H. R. Swapna
School of Commerce- PG Studies, JAIN (Deemed-to-Be University), Bangalore, India

S. Singh
Chandigarh University, N95, Chandigarh State Hwy, Ludhiana, Punjab 140413, India

D. Arpana
The Oxford College of Business Management, Bangalore, India

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_8

1 Introduction

Intelligent services become progressively fundamental with the development of the fourth step of Industry 4.0, where emerging technological innovations shift the fields of manufacturing as well as science [1]. Over the first three industrial revolutions, efficiency was improved by the introduction of modern mechanical, electrical, and computer innovations. In recent years, the desire to increase the standard of human life has contributed to further realization and more personal and interactive services development models [2].

The main consequence of the modern fourth industrial revolution is the development of cyber-physical networks and their output. Cyber-physical networks build software that integrate physical assets with computer capabilities [3]. CPS developments include a broad variety of uses such as energy networks, transit structures, surgical facilities, gas delivery. In CPS, contact with physical structures takes place through networks and performs complex analyses as data are extracted [4].

The usage of CPS networking, the Internet, and sensors contributes to the Internet of Things concept [5]. The Internet of Things may be regarded as a facility that facilitates cyber-physical networks since Internet of Things systems depend on protocols of communication, where physical assets will link, transfer, and exchange information. Since the IoT model does not test data and knowledge systems, cyber-physical networks are built to allow dynamic analytics utilizing the IoT connectivity architecture from a single analytical hub, with knowledge derived from raw data to deliver physical asset management commands [4].

The Internet of Devices and the cyber-physical networks both have a strong foundation for developing a new area of science: The Internet of Robotic Things (IoRT). This modern paradigm also incorporated many improvements in various domains including many applications that operate in difficult conditions [1]. For example, IoRT systems may be used autonomously and remotely by manufacturing industries to execute demanding tasks like production, packing, welding, quality control management, etc. Furthermore, IoRT structures have extend to libraries, athletics, and entertainment beyond the sector. Their growth, though, is partly due to the need to build integrated structures for Industry 4.0, [2] digital and real realms are mixing. In this example, IoRT stands for the heart of robotically integrated IoT systems that incorporate cloud computing and networking for intricate responsibilities, enabling robots to share, network and capture numerous types of human and computer knowledge [6]. A robot, cloud, and IoT convergence scheme can be tracked as mentioned in Fig. 1.

In recent years, several reports have been published, but classifies Internet of Robotics Things systems according to intelligent structures [3]. The existing study made an attempt to highlight Internet of Robotics Things applications aligned with numerous smart realms in Industry 4.0, defining the new Internet of Robotics Things innovations, and explaining how in our society, Internet of Robotics Things systems can play a crucial role. The main goal is to explain the primary issues in the field to explain, where Internet of Robotics Things implementations need to be more

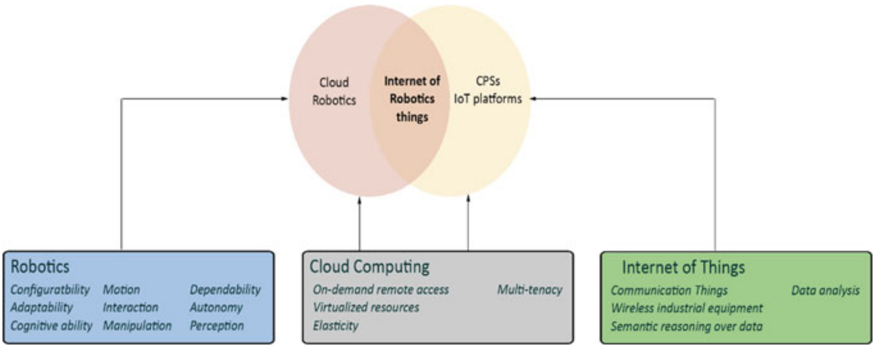


Fig. 1 Diagrammatic presentation of Internet of Robotic Things, robotics, and cloud computing

researched and applied [7]. As the objective of this manuscript is on manufacturing and processing, Internet of Robotics Things technology in smart growth and intelligent agriculture is thoroughly studied. Moreover, areas such as wellness, schooling, and monitoring are then discussed to explain how Internet of Robotics Things networks spread across several facets of daily life [8].

The manuscript is arranged accordingly. Section 2 displays the core characteristics of Industry 4.0, concentrating on CPS-based technology [9]. The Internet of Robotics Things is consequently introduced, and its elements and design are evaluated. The new robotic agents and applications in the health sector are described in Sect. 3 where IoT technology incorporates robotic structures. Section 4 addresses transparent obstacles mentioned in Sect. 5 draws the report's conclusions [10].

2 The Role of Smart Technologies in the Fourth Industrial Revolution (Industry 4.0)

Industry 4.0 is considered as the fourth industrial revolution and treated as the next manufactured stage, where digital convergence and smart innovation are used to change the link between equipment and humans. Industry 4.0's development has improved the whole manufacturing industry, establishing the so-called smart space and, in particular, smart factories [11, 12].

The smart computer idea is strictly related to the smart factory. These IoT-based systems have switches, add-on sensors and use their own modules in real time including other self-consciousness and autonomy machines [13]. Auto-consciousness allows machines predict potential malfunctions, and self-comparison encourages the computer to better customize the settings depending on its operating experience. For these reasons, all wired and wireless network designs and networking protocols help to link sensors and actuators for more specific industrial applications. Auto-consciousness allows machines predict potential malfunctions to ensure contact between different equipment and robots and ensure effective and productive manufacturing [14].

3 Framework of CPSs

Cyber-physical system is the most critical elements of Industry 4.0's. To further expand other scientific fields, cyber-physical networking technology becomes important as the Internet became accessible to all physical devices, integration of simulated and physical worlds for smart products and production. The central structure of the CPS is real universe, cyberspace and contact networks: (i) physical universe refers to the artifacts, systems, and ecosystems that are to be tracked or managed, (ii) cyberspace reflects such knowledge structures as utilities, software, and judgments, though (iii) communication networks are those elements that link cyberspace with true existence in the world [15, 16]. As the fourth industrial revolution begins, management and production processes have been very critical in many intelligent sectors. Smart goods and intelligent development structures are in particular narrowly linked to the CPPS and various architectures are meant to fix failures in discrete case processes [17].

CPPS will help boost the versatility of development processes focused on IoRT in smart fields, including production, livestock, medical operation and elderly care. In specific, human–robot interaction (HRI), which includes human beings and robots, in these areas just particularly in manufacturing [18]. In production processes, such as assemblage and welding the details generated by physical interaction between people and robots can be used. The idea of an architecture of the human–robot interaction is essential for the design and implementation of the cyber-physical distribution mechanism and seek to impact the direction of a robot dependent on human labor points [19]. Since the CPS is the first stage of growth, it is important to describe its framework and methodology. The CPS architecture is distinguished by five distinct levels in manufacturing environments that can be seen in Fig. 2. The so-called 5C architecture clarifies how a CPS can be built from initial data collection to final value output. However, many CPS architectures were created, each concentrating on different aspects [20].

In the sense of Industry 4.0, the goal is to better describe manufacturing structures and smart factories. The stages of 5C architecture are listed below.

- Internet of Robotic Things

The fourth industry innovation in recent years has culminated in the creation of the Internet of Robotic Things, that achieve control, vision and exploitation decision-making systems. Figure 3 displays a sketch of the IoRT modules.

The most sophisticated robotics principle is located in IoRT, where CPSs offer a clear base for developing IoT itself. New robotic developments have been paired with cloud computing in IoRT systems networking, CPSs and IoT protocols incorporation in the creation of emerging technologies. This modern method integrates

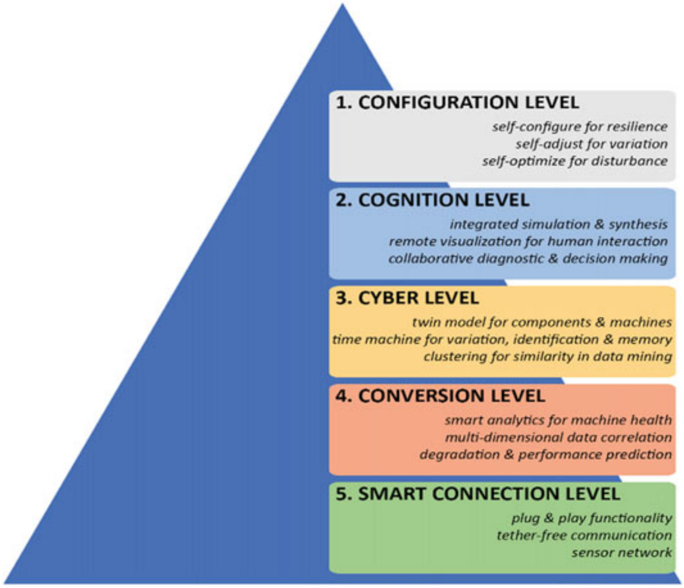


Fig. 2 Diagrammatic presentation of CPS architecture

multiple technology for complex activities and in a heterogeneous setting. In addition, wireless sensor networks (WSNs) have been an important topic in recent years. In particular, the need to mount sensor nodes according to such algorithms and circumstances implies that certain sensors may be built into robot networks [21]. These robots coexist with sensors to increase the WSN's capacity. The development of modern Internet of Robotic Things-related technologies and applications to solve different roles and software that can address diverse roles. Also, installation of robotic networks with wireless sensors was also allowed from deployment to connectivity [22].

Robotic devices have triggered major improvements in different areas of human life. Robots have been used in business and academia to conduct all kinds of complicated and difficult activities, such as packing, assembling, and welding. In this sense, the creation of modern heterogeneous robotic systems is a consequence of IoT and robotic convergence, in order to enhance the autonomous actions of robots. The incorporation of robotics and networking is also important for the creation of IoRT systems [23]. Networked robots combine device structures (both software and hardware) and network implementations to determine optimal intervention series of local and distributed details and then comply to change the physical state physically [24].

As shown in Figure 4, the core architecture of the Internet of Robotics is made of three main layers:

- Physical layer
- Network and control layer

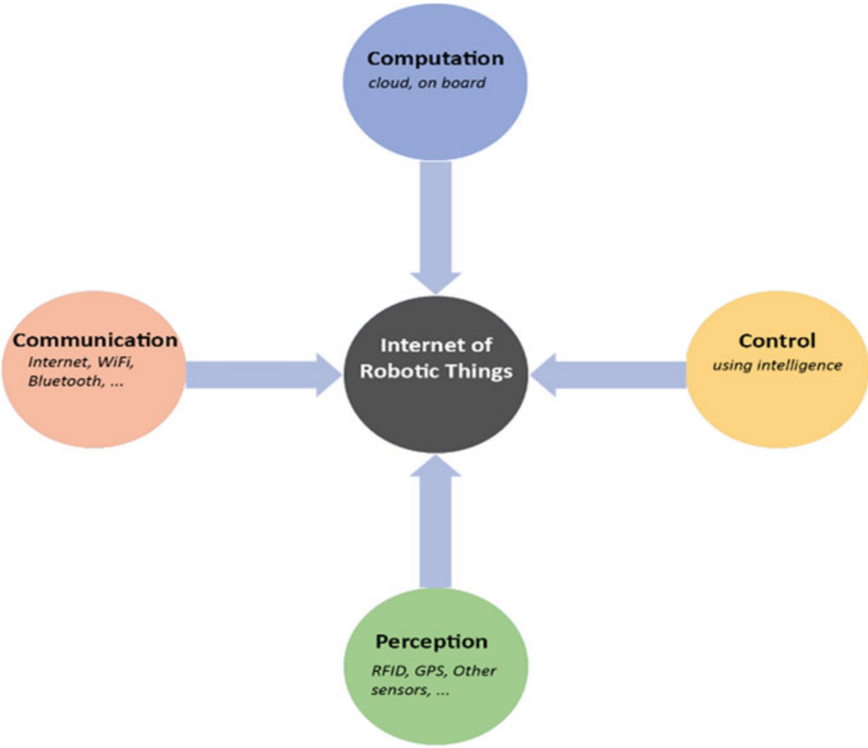


Fig. 3 Diagrammatic presentation of IORT architecture

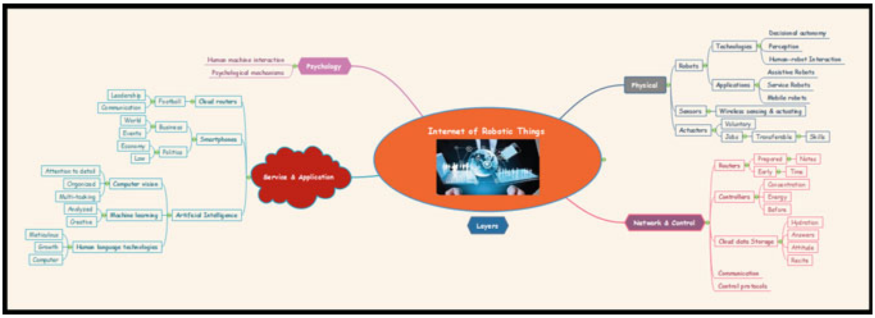


Fig. 4 Internet of Robotics Things architectures layers

- Service and application layer.

With robots, sensors and actuators, the physical and psychological layers are the lowest stage of IoRT architecture. In order to optimize, track and manage different procedures, such as robotic applications may also involve navigation, setup, tuning, sensors, and actuators. The total and efficient integration of sensors and actuators into the network and control layer for robot implementations is therefore feasible, where separate components can be interacted and managed by similar protocols [25]. Basic connection is used for smooth transfer of information between short- and long-distance robotic systems close field contact and the WSN.

Lastly, the layer of service and device determines the maximum standard of IoRT architecture, namely the deployment of programs in smart settings, to monitor, to process and to evaluate all environmental parameters (robots, sensors, and actuators). Additionally, this layer also contains artificial intelligence and machine learning algorithms that enables smooth incorporation of robotic systems and IoT systems in order to find personalized solutions to real-physical problems [26].

4 The Role of IoRT in Different Domains

In the fields of health treatment, schooling, and monitoring, IoRT networks are known to have tremendous importance, depicts a summary of the literature for this portion. Specifically, in certain applications, the IoRT may have well-being, community, and economic advantages, for special needs groups, including mental disorders, stroke survivors, patients, amputees, etc. IoRT systems may also provide several benefits for other uses, such as patient monitoring, personnel and clinics, automated data storage, and sensing (Table 1).

To concentrate on the education area, robots need suitable and adaptive behaviors to establish and sustain sufficient social connections with individuals, in order to be utilizable in resources of help such as homework and education. In comparison, EDA reactions assessed in children will vary greatly from a typical adult reaction since children cannot respond to such desires like adults do. Under these cases, the IoRT technology used for child education problems would capture, store, and interpret data such that child behavior and condition are automatically predicted.

It is a recognized problem to track areas and individuals in order to regulate human well-being and preserve certain habitats. IoRT devices have a vital role to play in delivering intelligent high-monitoring technology in locations including vulnerable places, clinics, military boundaries, government facilities, and households. Closed circuit TV cameras (CCTV) are typically used for indoor and outdoor tracking. However, such equipment needs certain obstacles and restrictions, mostly because of collateral damage and blind spots. The rise in the number of cameras inside the device will partially overcome those limitations and to cover more nooks while growing expense and sophistication of systems.

The addition and development of IoRT management tools tend to be the most powerful way to track contexts, while IoRT systems can easily and efficiently be designed and distributed to cover broader areas to defend a certain room. In addition, robotics-integrated cloud systems to ensure real-time tracking are also useful for remotely controlling settings such as buildings, factories, retail and wholesale shops, as well as the identification of human activity in multiple scenarios. Measuring

Table 1 Comprehensive review of role of IoRT in different domains

References	Domain	Technology	Findings
[26]	Agriculture domain	Robot navigation	Various UAVs are used for data processing through ground surveillance and mapping to varying levels of fertilizer, spraying, etc.
		Multi-robots	
		Path planning	
		Cloud computing	
[25]	Agriculture domain	Robot navigation	Cooperation between heterogeneous robots from the agricultural field utilizing a new method focused on a theory of independent events (DES) and Rama-Wonham (RW) that governs the complex dynamics of heterogeneous multi-robot systems used for intelligent agricultural applications
		Path planning	
		Multi-robots	
[5]	Health care (medicine)	Robot navigation	Indoor robot (CIoT) focused on the multimedia platform cloud and IoT, contact protocol RFID and IEEE802.11 and cloud platforms help medicine
		Cloud computing	
[22]	Healthcare system	Robot navigation	IoT-compatible telerobotic architecture designed to represent home-telebot-centric health care infrastructure, and combine the unit of robot operation with a human movement capture device. The used robot is a mutual double-armed robot named Yu Mi, imitating human behavior captured using wearable gesture selection systems
		Cloud computing	
		Human–robot interaction	
[23]	Healthcare system	Robot navigation	A health evaluation kiosk is introduced by the implementation of a robotic interface that ensures links to intelligent urban knowledge, connectivity networks and specialized tasks through designing apps that address patient needs
		Human–robot interaction	
[15]	Manufacturing domain	Robot navigation	A health evaluation kiosk is introduced by the implementation of a robotic interface that guarantees its availability in the smart city knowledge and connectivity networks and can provide unique tasks through designing apps that address patient needs
		Cloud computing	
		Human–robot interaction	
[27]	Manufacturing domain	Robot navigation	Error transmission based on the next iterative point algorithm for the robot manipulator to measure data from cloud points on different stereo viewing systems
		Path planning	
		Data gathering	

(continued)

Table 1 (continued)

References	Domain	Technology	Findings
[14]	Manufacturing domain	Robot navigation	The incorporation of robotics into CPPS in order to handle various weight items, integrating UGVs with robot manipulator and air moving systems for intelligent processing and smart production
		Cloud computing	
[19]	Surveillance domain	Robot navigation	IoT-based robot device, identified as the Inter Bot 1.0, fitted both with contact networks for long and short distances. The robot monitors smart monitoring in real-time settings
		Path planning	
[7]	Surveillance domain	Robot navigation	Surveillance robot for horizontal and vertical surfaces during the monitoring of surface shifts, explore room and send live video to the remote workstation through Wi-Fi
		Data gathering	
[8]	Surveillance domain	Cloud computing	Autonomous networked robots (ANR) are built on a WSN to control any sensor node, including smoke, infraround fire, odor, activity sensors and network touch and RF transceiver
		Robot navigation	
		Path planning	
[20]	Surveillance domain	Robot navigation	Multi-robot framework for monitoring and rescue activities focused on swarm intelligence with IoT real-time uploading data into the cloud, leveraging the wireless links between many officers, PID strategies and the ACO such that they can execute tasks synchronously
		Path planning	
		Data gathering	
		Multi-robots	
[21]	Education domain	Path planning	Architecture and architecture of a portable, cognitive processing, affective robot dubbed Fitbot. Such a robot may conduct the interpretation of multimodal data to understand the patient's emotions
		Cloud computing	
		Human–robot interaction	
[17]	Education domain	Robot navigation	Explore room and send live video to the remote workstation through Wi-Fi to episodes of focus, upheaval and enthusiasm in the skin
		Data gathering	

sensors including GPS, magnet area, air quality, and environmental values can also help to monitor indoor and outdoor situations when the robot passes data in real time during a monitoring mission.

5 Issues and Challenges

Internet of Robotics is a recent idea aiming to outline the integration of robots and IoT systems with cloud computing. The consequence of this convergence is an increasingly increasing relationship between IoT, cloud computing, and robotic science. The incorporation of smarter robot systems guarantees interoperability, creative technology, and implementations in real time and autonomous cooperation. However, certain IoRT implementations are not sufficient to be entirely translated to business scenarios, as the key studies in the subsequent literature are performed in academic laboratories.

A core problem in smart environments is the need for collaboration between different robots and the sharing of room with people. More than one challenge also remains for multi-robot collaboration with respect to consensus networks, control management, cooperation both against infrastructures and other robots and coordinated monitoring. The absence of help for heterogeneous robot configurations would also be a specific problem. In multi-robot activities, IoRT technology from various vendors is considerably challenging to combine, customize, and teamwork, sometimes utilizing different technologies. Several developments have been achieved for some functionality of multi-robot systems in the area of science, yet further advances ultimately include the robot producers' themselves' active involvement. Moreover, communications with human machines are becoming gradually pertinent in diverse locations, such as prisons, restaurants, and places of service. Smart robots will react to popular human gestures in HRI. The related data processing often strives to achieve full autonomy and stability in HRI. However, several restrictions continue to impede the production of HRI. New HRI forms, such as eye-tracking, speech interactions, and biodegradable recognition have been researched in recent years yet they still require testing, because several were just analyzed and have not yet been widely deployed in research laboratories.

As CPS and Business 4.0 develop remote operation innovations are being adopted by an increasingly lucrative sector, where human operators operate remotely, reprogram, and control robots from a protected space. In certain circumstances, HRI plays a new and valuable role in the need to improve the technology, and it can be highly beneficial for remote supervision of manufacturing robotics by human operators. Advances in IoRT technology would respond to a faster, more modern form of handling industrial activities. Just, remote study has drawn tremendous attention owing to its immense advantages. Remote job was found to boost production quality due to an improved balance between work and life. IoRT systems will further accelerate those developments, opening up different possibilities for remote usage of industrial robotics. Moreover, every other implementation of this survey will conveniently export the same paradigm. In school, for example, in an IoRT scenario, students learn by incorporating social robots. These problems contribute to deepening of connections between human beings and robotics for schooling. In addition, more research on IoRT technologies would be of extreme importance. Besides improving multi-robot and HRI systems in smart settings, the development of IoRT applications is a major concern in the age of Industry 4.0. In recent years, it has the interest in examining energy spending among researchers been growing. The challenge in measuring and enhancing energy quality in a smart world derives primarily from an inadequate knowledge of the behavior of energy use. This expertise, for example, needs to be integrated into efficient manufacturing management to achieve successful processes on a long-term basis.

In these applications, additional computation and saving features should be applied to the IoRT network, since robots cannot handle and retain vast amounts of data at the edge. The major challenges with cyber protection include insecure user and robot connectivity, authentication difficulties, confidential data leakage, and bad default robot setup.

As protection in networked robotics is fundamental, modern smart network architectures become necessary to secure not only data knowledge, but also individuals in HRI systems. Indeed, human–robot protection cooperation is a crucial concern in the automotive world, mainly for IoRT systems. The smooth running of such industrial networks will jeopardize cyber-attacks from the network or through the Internet. To address these issues, CPS protection vendors need to help to distinguish possible threats by comparing, data processing and compilation from different sources. Taking into consideration, the period expended on person and machine designing IoRT structures, classification and near monitoring of production data and other industrial management applications, since it might be necessary to properly collect information background in order to maintain cyber-physical protection. The output data should be maintained in the same format with the same material as it was produced and should be shared with any other compatible industrial system worldwide utilizing the same protected protocol.

Through unified verification and authorization procedures, cyber-physical protection challenges can be solved by streaming output details such as data pieces, assigned liability, managed operator access and granular licensing. Since the connection between intelligent devices is critical for further development in business and research, solving cyber-physical security problems is essential to boost IoRT systems' growth in smart spaces. Protection concerns with both CPS and robot links must also be studied further.

6 Conclusions

The manuscript highlighted the development of primary Industries 4.0 that helped develop IoRT structures. A description of how CPS handles the relation between actual and virtual reality to incorporate an internet importance. The architectures of both systems have been identified and numerous IoRT implementations in various intelligent areas mentioned, showing how IoRT-based devices carry out conventional robotics. The convergence of robotics with the IoT and artificial intelligence would be critical. IoT provides the ability to communicate with various stakeholders such as applications, smartphones, and people contact, providing the right option for multiple application domains. The combination of robotics, IoT and artificial intelligence results in robots that can execute more complicated tasks independently or in collaboration with humans.

In fact, the state-of-the-art technologies in smart realms is evaluated after 2018, revelation that IoRT devices are now central in multiple contexts, beginning with industrial ones, such as manufacturing and forestry, and impacting regular life, for example, health care, education sector, and surveillance. IoRT applications will also pave the groundwork for the growth of other sectors beyond the manufacturing sector, such as culture, museums visit, and athletic events, to enhance progressively the facets of human existence. Furthermore, it has been noted that designing IoRT systems will provide an effective solution to the need for remote work, through which the current criteria of remote human–robotic interactions can be the key to greater fulfillment and efficiency.

The key subject of this study was those IoRT-based structures in the manufacturing and development sectors. IoRT problems and obstacles have been extensively examined and highlight. The increasing need to study in the partnership between robots and robots, particularly in heterogeneous robotics, human–robotic interfaces for oriented human communications, energy management for performance optimization and computer defense for critical data safety. The expected increases in performance, robustness and protection would open up new doors for even new applications that can benefit from new developments in the fields of robotics and networking.

References

1. Stock T, Seliger G (2016) Opportunities of sustainable manufacturing in industry 4.0. Procedia CIRP 40(Icc):536–541
2. Marvel JA, Bostelman R, Falco J (2018) Multi-robot assembly strategies and metrics. ACM Comput Surv 51(1):1–44
3. Marvel JA, Bostelman R, Falco J (2018) Multi-robot assembly strategies and metrics. ACM Comput Surv 51(1)
4. Architectures R (2018) A comprehensive survey of recent trends in cloud
5. Romeo L, Petitti A, Marani R (2020) Internet of robotic things in smart domains: applications and challenges. Sensors 20(3355):2–23
6. Lee YK, Goh YH, Tew Y (2018) Cyber physical autonomous mobile robot (CPAMR) framework in the context of industry 4.0. MATEC web conference vol 167
7. Lu Y (2017) Industry 4.0: a survey on technologies, applications and open research issues. J Ind Inf Integr 6:1–10
8. Kehoe B, Patil S, Abbeel P, Goldberg K (2015) A survey of research on cloud robotics and automation. IEEE Trans Autom Sci Eng 12(2):398–409
9. Chen S, Li Y, Kwok NM (2011) Active vision in robotic systems: a survey of recent developments 30(11)

10. Cecil J, Albuhamood S, Ramanathan P, Gupta A (2019) (Smart CPS) An internet-of-things (IoT) based cyber manufacturing framework for the assembly of microdevices. Int J Comput Integr Manuf 32(4–5):430–440
11. Xue Y (2018) Mobile robot path planning with a non-dominated sorting genetic algorithm. Appl Sci 8(11)
12. De S, Zhou Y, Abad IL, Moessner K (2017) Cyber-physical-social frameworks for urban big data systems: a survey. Appl Sci 7(10):1–26
13. Simoens P, Dragone M, Saffiotti A (2018) The internet of robotic things: a review of concept, added value and applications. Int J Adv Robot Syst 15(1):1–11
14. Yao X, Zhou J, Lin Y, Li Y, Yu H, Liu Y (2019) Smart manufacturing based on cyber-physical systems and beyond. J Intell Manuf 30(8):2805–2817
15. Kumar A (2018) Methods and materials for smart manufacturing: additive manufacturing, internet of things, flexible sensors and soft robotics. Manuf Lett 15(12):122–125
16. Md Fauadi MHF, et al (2018) Intelligent vision-based navigation system for mobile robot: a technological review. Period Eng Nat Sci 6(2):47–57
17. Ahmad Yousef KM, AlMajali A, Ghalyon SA, Dweik W, Mohd BJ (2018) Analyzing cyber-physical threats on robotic platforms. Sensors (Switzerland) 18(5):1–22
18. Ghosh P, Gasparri A, Jin J, Krishnamachari B (2019) Robotic wireless sensor networks 164(5)
19. Dixit S, Dhayagonde M (2014) Design and Implementation of e-surveillance robot for video monitoring and living body detection. Ijsrp Org 4(4):1–3
20. Shakeri R et al (2019) Design challenges of multi-UAV systems in cyber-physical applications: a comprehensive survey and future directions. IEEE Commun Surv Tutorials 21(4):3340–3385
21. Chen M, Zhou J, Tao G, Yang J, Hu L (2018) Wearable affective robot. IEEE Access 6:64766–64776
22. Zhou H, Yang G, Lv H, Huang X, Yang H, Pang Z (2020) IoT-enabled dual-arm motion capture and mapping for telerobotics in home care. IEEE J Biomed Heal Inf 24(6):1541–1549
23. Romeo L, Petitti A, Marani R, Milella A (2020) Internet of robotic things in smart domains: applications and challenges. Sens (Switzerland) 20(12):1–23
24. Gaur K, Virdi GS (2018) Smart manufacturing with robotic automation and industry 4.0-a review. Int J Innov Res Sci 7:2481–2486
25. Ju C, Il Son H (2020) Modeling and control of heterogeneous agricultural field robots based on Ramadge-Wonham theory. IEEE Robot Autom Lett 5(1):48–55
26. Tsouros DC, Bibi S, Sarigiannidis PG (2019) A review on UAV-based applications for precision agriculture. Information 10(11)
27. Pomerleau F, Colas F, Siegwart R (2015) A review of point cloud registration algorithms for mobile robotics. A Rev Point Cloud Regist. Algorithms Mob Robot

Security Issues and Vulnerabilities in Web Application

Sitara Anumotu, Kushagra Jha, Amit Balhara, and Pronika Chawla

Abstract Role of web security has become an important topic as web and web application became quite demanding and people started using them every day. Last year we saw slight growth in web attacks and exploitation of their vulnerabilities. For example: Recently on the website of one of the largest airlines, Air India, an attack was performed where user's credit card and sensitive information was captured. Input validation issue has resulted in many web application vulnerabilities. Broken access control, improper error handling, etc., are some of the examples of web security vulnerabilities. It is quite easy to acknowledge and eschew many of the web vulnerabilities. Regrettably many of the web developers are not security aware, which causes many of the web pages and websites to be vulnerable on the Internet. In this paper, Authors have discussed some popular web security vulnerabilities to detect them by providing some security mechanisms and some results to safeguard our web applications. It is a challenging and difficult task to protect our web application from vulnerabilities as nearly every day new techniques, tools, and methods are being invented by the attackers. It is one of the essential parts to know what various types of vulnerabilities are and how to detect and secure them from attackers.

Keywords Vulnerability · Vulnerability detection · Scanner · Security · Web application

1 Introduction

In this world, the World Wide Web has become one of the mediums for communication. To send messages, watch videos, play games, gaining knowledge, making transactions, shopping, and for many more different purposes millions of people

S. Anumotu (✉) · K. Jha · A. Balhara · P. Chawla
Department of Computer Science and Engineering, Manav Rachna International Institute of Research and Studies, Faridabad, Haryana, India
e-mail: sitara.a.952@gmail.com

P. Chawla
e-mail: pronikasoft@mriu.edu.in

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_9

connect to web application every day. Some of the users try to exploit vulnerabilities of these critical web pages and websites, which can lead data breaching. Some of the popular vulnerabilities are bad input, broken access, broken authentication and sessions management, improper error handling, parameter modification, insecure configuration, cookie modification, and directory traversal. Consequences can be disastrous if the attacker manages to exploit the vulnerabilities. It might compromise the information of the user and the company or organization.

2 Literature Review

Vieira, M.; Antunes, N.; Madeira, H. in their article "Using web security scanners to detect vulnerabilities in web services" studied about various scanners available to detect and scan the vulnerabilities in a web application and compare their execution which will fit to put in a web service [1].

Silic, Marin; Krolo, Jakov; Delac, Goran stated in their article "Security vulnerabilities in modern web browser architecture." They studied about architecture of web browser as foremost origin of a vulnerability in a browser [2].

According to Fonseca J., Vieira, M.; Madeira, H. "Testing and comparing web vulnerability scanning tools for SQL injection and XSS attacks" on web application to identify vulnerabilities, automatic vulnerability scanner is used [3].

According to Johari, R.;Sharma, P., In the "A survey on Web application vulnerabilities (SQLIA, XSS) exploitation and security engine for SQL injection," the various types of structured query language injection attacks and vulnerabilities and security technique of XSS is the for most objective of this paper. On structured query language injection attack, they proposed a future study [4].

After studying various papers and articles on security and vulnerability of web applications, they all focused mainly on different tester, detectors, tools, and scanners but did not highlight any prevention techniques or how to protect web application against vulnerability. In this paper, we not only discussed about some popular vulnerability but also proposed some methods to detect and protect web application by providing some security mechanism.

To test whether a web application contains bugs and vulnerabilities, there are mainly two approaches:

- White box testing: In white box testing, to find or track down any vulnerability or faulty code line, source code of the application is inspected. By creating add-on tools for common development environments, this operation can be often integrated into the development process.
- Black box testing: In black box testing, inspection of the source code is not done directly. Instead, to the web application, special input test cases are generated and sent. It indicates errors or vulnerabilities by inspecting for unexpected behavior by the results returned by the application.

For finding bugs and vulnerability, white box testing has not accomplished widespread use. Bounded detection capacity of white box analysis tool, heterogeneous programming environment and the complexity of web application are some important reasons for the limitation of the white box testing. To discover security problems in web application, black box vulnerability scanner is used in practice.

So, through black box testing of each vulnerability, we demonstrated the attack by building blog a website and exhibited by secure and insecure version of the website [5–7].

3 Web Application in Different Sector

Web applications have become an important part of our lives, as they help us to complete most of our everyday activities. Here are some examples of web applications deployed in different sectors.

3.1 Health Care

So, many patients have started using healthcare apps, as these apps are increasing in the healthcare sector. Now patients are playing a role of a consumer in this telehealth market. We can see the growth of telehealth market in recent years. By using artificial intelligence, web and mobile application technology, healthcare companies have found easy solutions to interact with the patients virtually and get to know their health issues. Application helps patients to buy medicines online and deliver it to them. Apart from other sectors, healthcare sector occupies more time and effort to record a large amount of condemning data manually. Different functions like monitoring, communication, sharing data, reminding, and giving clinical support are equipped in these apps [4, 8].

3.2 Travel

Travel apps are very easy to use and accessible. To travel anywhere across the world, traveler depends on the web travel apps. Through the app, the traveler can book flights, hotel rooms, and so on. Traveler can now access details about flights and can track their location and arrival and departure time, information regarding gate changes, and many more through the latest travel apps available.

3.3 Educational

The image of education has been changed by educational web applications. Earlier, students studied from their old academic books, attended long lectures, and gave examinations. But now, a student can access anything through these educational applications like self-testing and understanding concepts very well as web applications have games and quizzes which makes it more interactive and fun. Now, schools are also introducing these apps as it analyzes data, and it can acknowledge each student's area of interest and can help students to improve.

3.4 Web Application for Home

Day-by-day popularity of smart homes is increasing. By 2025, IOT-based applications in the home will have impact of $200 billion to $350 billion. You can control each corner of your house using these smart homes with a single interface remotely. There is voice command recognition which can help people with disabilities and help them to have the same quality of life like a normal person [9–12].

4 Security Requirements

1. For security vulnerabilities, the web application should be tested and reviewed properly.
2. Authorized individuals should approve and be coordinated by vulnerability assessment.
3. All security defects should be classified by seriousness. All defects must be recorded in a defect tracking system, which is properly recognized as a security vulnerability. Before the application is released, this information must be properly secure and be fixed.
4. Prior to production deployment, methods for correcting security problems must be specified and documented.
5. Developer should use OWASP or server coding to develop web software application.
6. At least with one individual or peer member who is technically trained, should review the code.
7. Data received through HTTP request should be validated. Cross site scripting, SQL injection, and HTTP responses are some attacks which raises when data is not validated.
8. Server-side data should be validated. Malicious user can modify all the data, so there should be validated server side.

9. Whenever anything in the application is to be changed or produced, vulnerability scan should be performed.
10. Via SSL cookies and session IDs should be passed. User's session can be compromised by unprotected session IDs and cookies and system security.

5 Vulnerabilities in Web Application

We have discussed some vulnerabilities which are independent of the primary platform, and which can be commonly found in mercantile sites. These vulnerabilities are quite popular and are known for a long time. Web servers and web application environments are distrusted for these vulnerabilities.

5.1 Injection Flaws

Failure in filtering untrusted input results in injection flaws. It can occur when unfiltered data is passed to the SQL server (SQL injection), the browser, the LDAP server (LDAP injection), or anywhere else. The issue here is that the attacker can inject commands into these entities, resulting in data loss and browser takeover [13]. Any data received from untrusted sources must be screened, preferably using a whitelist. A blacklist should almost never be used, as getting it right is difficult and usually easy to circumvent. Antivirus software solutions are frequently used to demonstrate how blacklists fail. Matching patterns does not work.

5.2 Broken Access Control

Categorization of the user is done by the different level of privileges in web application. How the web application should authorize access to functions to some users and not others is determined by the access control. It is also known as authorization. But sometimes may be higher level of authority can be accessible by the attacker.

5.3 Insecure Configuration

Someone unauthorized should not be accessible to web configuration files and directories as the web server that hosts the web application is incorporated by web configuration files and directories. The attacker should be kept away from getting access to these files and directories.

5.4 Error Handling

When in a web application, under normal transactions if error occurs and not handled by sending an error message to the user, attacker might get a clue about the flaunt in the application and can disturb the users. • Invalid inputs are requested from the user to control or regulate the response to provide the web server by web application. To the web application, request input values are entered by the user. To detour the websites security, attacker may insert some damaging information to the web application.

5.5 Broken Authentication and Session Management

When the user is logged in, session is created by the web application. The interaction period the user interacts with web application is specified. To maintain state, user is provided with a distinctive or unique Id by using session. The Id used between the user browser and web server is stored in a cookie. Attacker can exploit and steal it if session details are not protected or defended properly.

5.6 Parameter Modification

When the attacker does not fill the form but passes parameter from the URL itself then problem occurs in parameter modification. Overall site data and form data can be affected (Table 1).

6 Security Solutions

1. Take help from a professional and ask them to attack on your web application. By this way, you will get to know the vulnerabilities your web application consists.
2. Keep yourself updated about web security and new attacks. You can follow some security blogs to keep your knowledge running.
3. Always backup your data. If your application is subjected to any data breach or any malware enters your system, you can easily get all your data restored.
4. Always keep scanning your application for vulnerabilities, so that your application can be safe, and no major loss would happen.
5. You should hire a security expert, as they can monitor and keep track of all new malware and attacks.
6. User output should be sanitized.
7. Everything should be up-to-date. Make sure your application is updated.

Table 1 Security vulnerabilities and their preventions in tabular form

Web security vulnerabilities	Preventions
Injection flaws	All the inputs from the user are needed to be filtered by the firewall before entering the system
Broken authentication	Admin should at least change default password and enable two-factor authentication
Cross site scripting	Website should not return any HTML tags to the client
Unsure direct object references	Website should perform proper user authorization and use of session variables and help
Security misconfiguration	Proper security tests should he performed before launching the website
Unauthorized data exposure	Sensitive data should be always encrypted
Lost function level access control	Proper authorization should be performed on the server side
Cross site request forgery	Enable password for any modification in website and verify a secret token which is always hidden from any 3rd party site or application
Using components with known loopholes	Every code or application should be kept up-to-date and known errors should be resolved as soon as possible
Unvalidated redirects	Use minimum redirects

8. Strong passwords should be used. Inserting symbols and numbers makes it hard for the attacker to get access.
9. In addition to strong passwords, two-way authentication is the best way to protect your application. In two-way authentication you will get a mail or an OTP before signing in. Your application will be safe.

7 Result Analysis

Following are the results of some of the vulnerabilities:

7.1 Broken Access Control

Attacker may try to access the unauthorized page, redirect the page to custom error page message (Figs. 1 and 2).

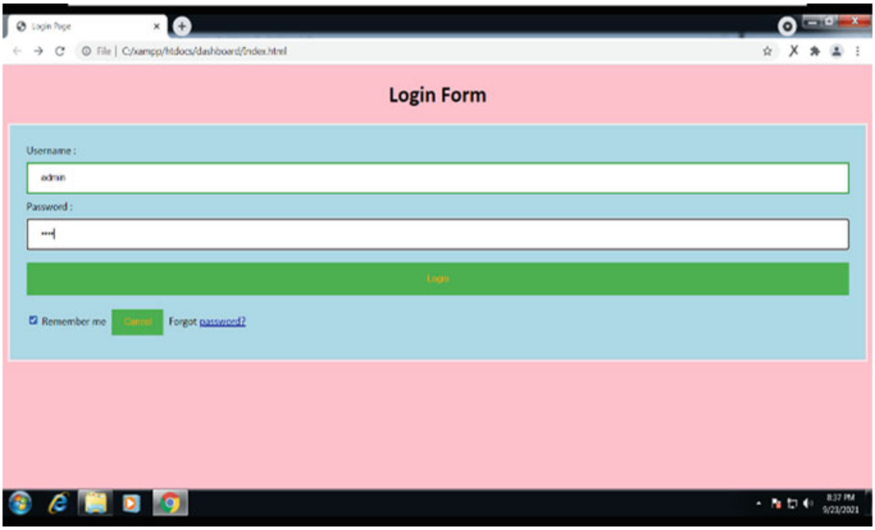


Fig. 1 Broken access control result

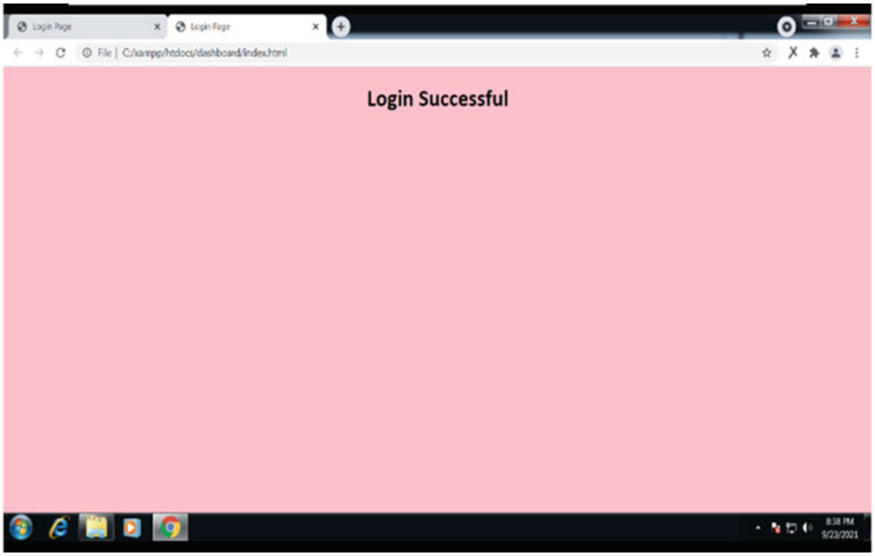


Fig. 2 Broken access result

7.2 Session Management

Open the login page, fill the form, and login if session handling is not provided. Error will be generated without session and access cannot be given (Fig. 3).

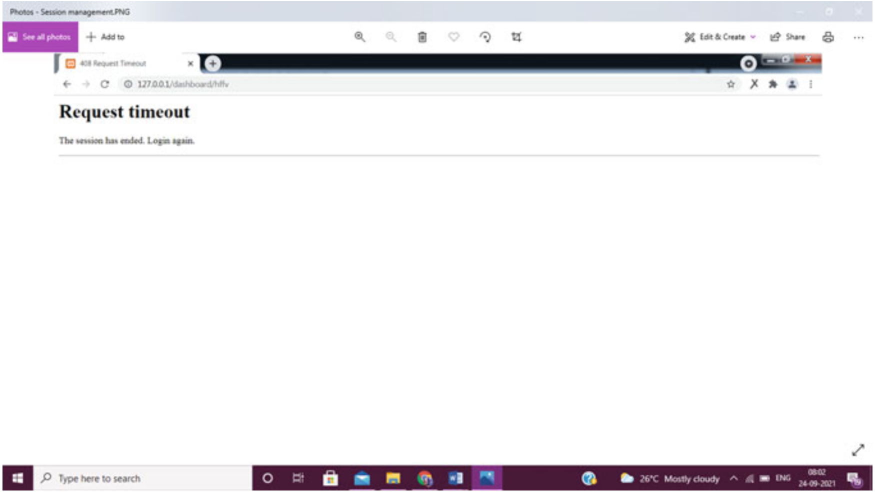


Fig. 3 Session management result

7.3 *Improper Error Handling*

Specific designed result should be provided when error occur which is helpful to the user without revealing internal details. Attacker will be redirected to custom error pages if they pass non existing URL (Fig. 4).

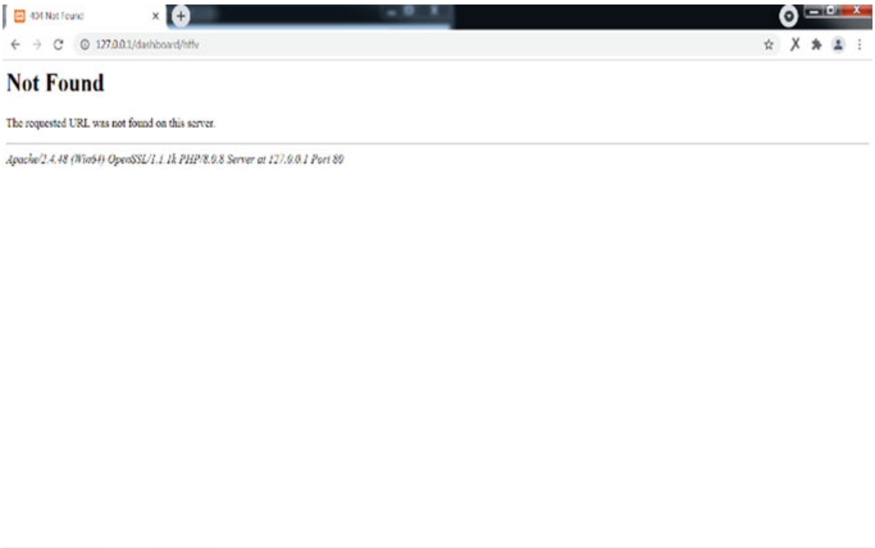


Fig. 4 Improper error handling result

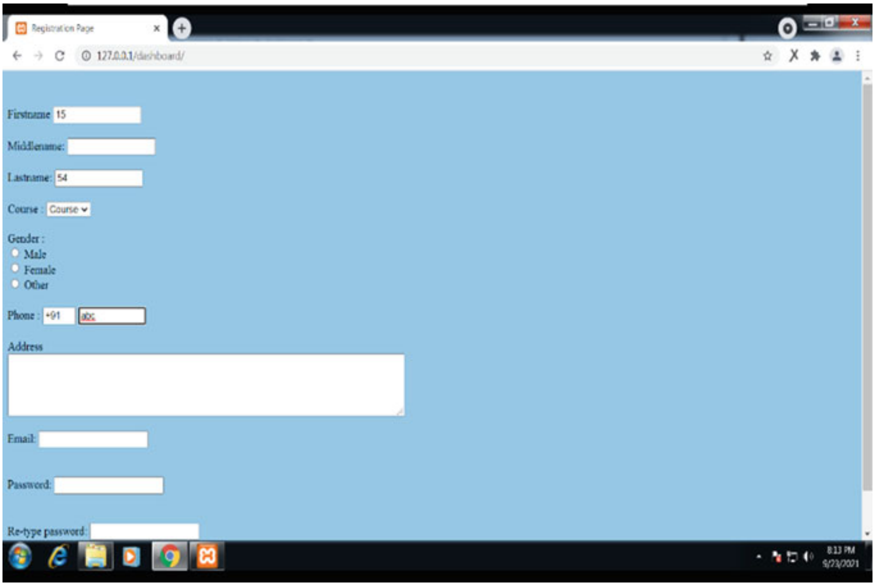


Fig. 5 Invalid input result

7.4 *Invalid Input*

Checking all input web applications against a strict format that specifies what input is valid (Fig. 5).

7.5 *Parameter Modification*

In this, attacker cannot login cause web page does not allow because attacker tries to pass parameter in the URL of the web application (Fig. 6).

8 Conclusion and Future Scope

As discussed throughout, security is among the most vital and neglected aspect when it comes to web application. As the web application boom continues to grow, application developers are faced with deadlines which are stricter than ever before leading to security taking a back seat. The demand web applications by newly minted businesses and organizations have outpaced web developers who are struggling to check all boxes in relation to security and vulnerabilities. In experienced application, development outlets often build use case apps which have security ramifications.

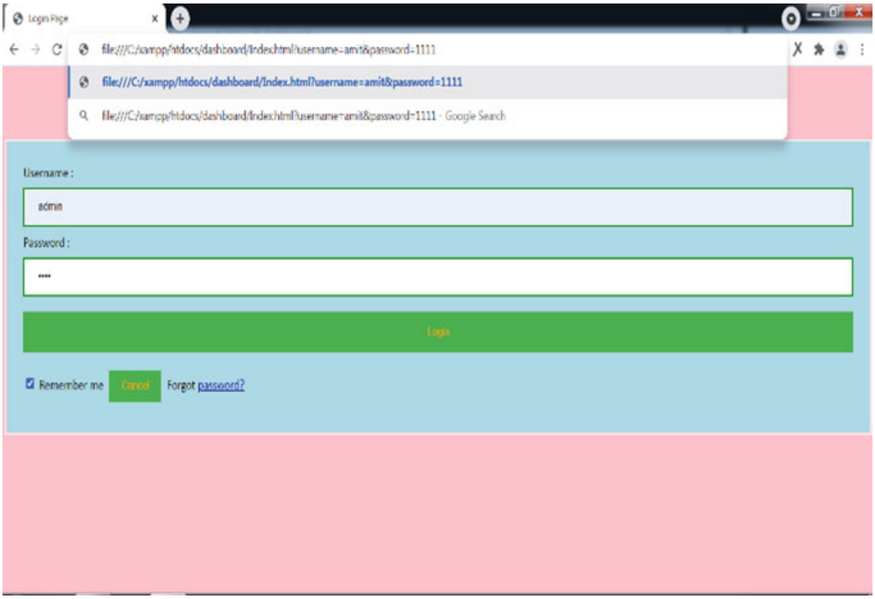


Fig. 6 Parameter modification

When writing code, developers and security professionals should have fundamental training and understanding, such as needing security aspects from the beginning in the requirements stage, such as input validation, error management, secure data handling, and other features. They should also incorporate security features into the software, such as adequate input validation to prevent command injection attacks on the application. This, together with automated security QA scanning and reporting techniques like "black box" and "white box," ensures that the programmer is secure even before it is released. Security of web applications can only be ensured by adopting practices such as effectively writing code and developing robust algorithms to avoid postproduction fixes. One might argue that firewall might be effective but developing an application firewall is not as feasible as it sounds and is particularly an itchy task to integrate alongside the existing application framework.

References

1. Vieira M, Antunes N, Madeira H (June 2009) Using web security scanners to detect vulnerabilities in web services. In: 2009 IEEE/IFIP international conference on dependable systems and networks. IEEE, pp 566–571; Rose K, et al (1915) The internet of things: an overview. The internet society (ISOC), vol 1, pp 1–50
2. Šilić M, Krolo J, Delač G (2010) Security vulnerabilities in modern web browser architecture. The 33rd international convention MIPRO, pp 1240–1245
3. Fonseca J, Vieira M, Madeira H (Dec 2007) Testing and comparing web vulnerability scanning

tools for SQL injection and XSS attacks. In: 13th Pacific Rim international symposium on dependable computing (PRDC 2007). IEEE, pp 365–372
4. Johari R, Sharma P (May 2012) A survey on web application vulnerabilities (SQLIA, XSS) exploitation and security engine for SQL injection. In: 2012 international conference on communication systems and network technologies. IEEE, pp 453–458
5. Shrivastava A, Choudhary S, Kumar A (Oct 2016) XSS vulnerability assessment and prevention in web application. In: 2016 2nd international conference on next generation computing technologies (NGCT). IEEE, pp 850–853
6. Idrissi SE, Berbiche N, Guerouate F, Shibi M (2017) Performance evaluation of web application security scanners for prevention and protection against vulnerabilities. Int J Appl Eng Res 12(21):11068–11076
7. Kindy DA, Pathan ASK (June 2011) A survey on SQL injection: vulnerabilities, attacks, and prevention techniques. In: 2011 IEEE 15th international symposium on consumer electronics (ISCE). IEEE, pp 468–471
8. Rexha B, Halili A, Rrmoku K, Imeraj D (Nov 2015) Impact of secure programming on web application vulnerabilities. In: 2015 IEEE international conference on computer graphics, vision and information security (CGVIS). IEEE, pp 61–66
9. Albishi S et al (2017) Challenges and solutions for applications and technologies in the internet of things. Procedia Comput Sci 124:608–614
10. Sedrati A, Mezrioui A (2018) A survey of security challenges in internet of things. Advan Sci, Technol Eng Syst J 3(1):274–280
11. Vignesh R, Samidurai A, et al (2017) Security on internet of things with challenges and countermeasures. IJEDR 5(1):417–423. ISSN 2321-9939
12. Alhalafi1 N, Veeraraghavan P (2019) Privacy and security challenges and solutions in IOT: a review. International conference on smart power and internet energy systems. IOP publishing
13. Internet of things best practices. https://docs.microsoft.com/en-us/azure/iot-fundamentals/iot-security-best-practices

A Systematic Review on Usability of mHealth Applications on Type 2 Diabetes Mellitus

Kamaldeep Gupta and Sharmistha Roy

Abstract In this present digital world, smart phones play a vital role, especially in the healthcare sector. The various mHealth applications installed in smart phones can solve primary health-related issues at a fingertip. However, the utility of those applications will increase if they meet user satisfaction and can solve the tasks effectively and efficiently. Thus, usability evaluation of mHealth applications is a matter of concern. Many usability evaluation techniques have been used till now but a majority of them are not unified and do not safeguard all usability aspects especially for mHealth applications. The purpose of this study is to recognize specific attributes that might help tremendously in assessing the usability of mHealth applications and examine the features of various usability evaluation methodologies for evaluating mHealth applications featured for diagnosing Type 2 diabetes mellitus (T2DM).

Keywords Usability · mHealth applications · Usability inspection and usability testing

1 Introduction

The availability and the popularity of mHealth applications have rapidly increased [1]. The intervention of mHealth applications has become an essential instrument of the public health professional and the researcher [2–4]. The development of mobile health applications for the management of the chronic diseases like diabetes is frequently taking place [5].

Type 2 diabetes mellitus (T2DM) is a kind of disease related to metabolic disorder. High levels of blood glucose are defining feature of this condition. Frequent urination, increased hunger, and thirst [6] are the common symptoms of T2DM. Self-management of diabetes involves actions that include condition handling and slowing down disease progress [7]. Self-monitoring of blood glucose (SMBG) aids in glycemic management, delays the onset of complications related to diabetes,

K. Gupta · S. Roy (✉)
Faculty of Computing and Information Technology, Usha Martin University, Ranchi, India
e-mail: sharmistha@umu.ac.in

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_10

and reduces the risk of hospitalizations [8]. Self-management of diabetes helps in improving one's lifestyle that includes proper exercise, a healthy diet, and adhering to medication [9].

Today, mobile applications are providing services and functionalities in the field of business, entertainment, and education and also in managing and preventing chronic diseases that include self-care of diabetes [10]. Assessment and evaluation of the usability of the mobile applications are very much important for ensuring the accuracy and utility of the mHealth apps that we use. The term "usability evaluation" refers to the process of assessing the mobile applications' usability. The evaluation of the usability of mHealth applications related to diabetes self-management is critical in assisting the patients who are diabetic in choosing the optimal mHealth application for their needs. The goal of this study is to examine and compare various usability evaluation methods (UEMs) for evaluating and assessing the usability of mHealth applications on Type 2 diabetes mellitus.

The comparison may serve as a reference that can guide the evaluators in choosing the most appropriate UEMs. The paper also aims to review the previous studies and the techniques through the systematic literature review (SLR) for the purpose of usability evaluation. This will help in selecting the usability guidelines.

This paper has been organized in the following way. Section 2 provides the related work on the usability evaluation of mHealth applications. The objective of this paper has been discussed in Sect. 3. Section 4 presents the usability attributes. Section 5 focuses on the results and analysis, and Sect. 6 concludes our work.

2 Review of Literature

Various usability evaluation methods are used to recognize usability issues and improve interface design usability. Analytical and empirical approaches are the two types of methods. Analytical methods are used for the inspection of the interface by involving usability experts. They are quick and cost-effective. On the other hand, in empirical methods, the testing involves the actual users. Table 1 shows the comparison study of various research works along with result analysis, advantages, and disadvantages. The focus of this section is on previous studies that assessed or evaluated the usability of various mHealth applications using various usability variables such as users, assessors, or software tools.

Table 1 Comparison study of usability evaluation methodologies adopted in various research works

Objective of the paper	Methodology used	Result analysis	Advantages	Limitations
The goal of this study [12] is to evaluate the usability of four diabetes applications that are intended to help patients manage their diabetes with the help of Self- Determination Theory (SDT), a health behavior theory linked to motivation and personality, for usability assessment and application design	Heuristic evaluation method: On the basis of Nielson's ten heuristics and usability satisfaction rating, experts evaluated four of the best-rated diabetes apps. Experts used a checklist for intuitive design, which was then updated for diabetes-related applications. Expert evaluators assessed the usability of seven device features, including carb consumption, exercise operation, insulin dosage, blood glucose (BG) reading, BG report, and BG report emailing. Minor, moderate, severe, and catastrophic violations have been assigned a severity rating (1–4). The SUS (System Usability Scale) was used to assess the app's usability and estimate the user's satisfaction level	• None of the four applications acquired a SUS rating related to the usability of 70 or higher. The average usability evaluation time of the application was 52 min • Total 314 heuristic evaluations were identified. 62% of the cases were extremely serious, with major or catastrophic severity ratings	• Evaluation of the four top-rated diabetes applications can improve usability • Intuitive design will help in making the applications patient-friendly. The application developers and designers can help to motivate patients and also provide education in better self-management of diabetes	• The future research should include the needs of patient training, usage barriers, and need of facilitators for use on a long-term basis • Inputs from the caregivers and the clinicians should be taken into consideration
The assessment technique in this work [13], involves the collection of usability data using a hybrid methodology and heuristic evaluation, as well as a thorough investigation of a mHealth approach for diabetes self-management	Heuristic evaluation, think-aloud, and questionnaire methods: A group of T1 diabetes patients who were using the Diabetes Monitoring (DiMo) application underwent a usability evaluation. To identify the usability issue, a hybrid strategy method combining think-aloud and post-questionnaire procedures was used. For data analysis, such as recognition, grouping, and structuring, Framework Analysis (FA), and Usability Problem Taxonomy (UPT) were used	• Each of the DiMo applications had a total of 28 consolidated issues. The Travis Criteria was then used to categorize these issues with the severity level • Study presented can be used for determining the usability level of the software applications	Usability evaluation results of the experiment can improve the tools of DiMo thereby improving essential factors such as empowerment and engagement in the monitoring of diabetes	Currently, these types of applications are rapidly expanding, and usability issues have a negative effect on device users. As a result, increased usability can lead to increased user empowerment and involvement

(continued)

Table 1 (continued)

Objective of the paper	Methodology used	Result analysis	Advantages	Limitations
The approaches for evaluating or assessing the content, usability, and efficiency of commercially available mHealth applications are summarized by the authors in [14]. The application features are compared using the Content Analysis	Content analyses and usability testing methods: • Analysis of the content, observational studies, usability testing, and efficiency or efficacy testing are used • Coding and interpretation of qualitative text-based content were employed in content analysis. Content analysis is done of the features, health information, and advice of the commercial health applications • Usability testing is done to see how an application functions and whether the application serves its intended purpose or not • Observational studies are done to assess the usage of the application, satisfaction, and the analytical value of application use • Efficacy testing determines a significant change in behavior and clinical results	• Three frequent comparators against which the contents of the application can be compared include clinical care guidelines, procedures based on evidence, and change in behavior or modification strategies • The application's content has been compared with clinical guidelines • Evidence-based protocols have been reported to have poor rates of strategies • Behavior change techniques involve theories related to health behavior and meta-analyses	• The application features using clinical guidelines behavior change and evidence-based protocols techniques can be compared • Usability testing can help to figure out how well the application functions and achieves its aim	• Systematic review of the applications was not conducted • Researchers should minimize delays in analysis of data, processing of data, and publication due to the rapid rate of technological progress • Evaluation of the technical functions of the application is not done
The authors of this study [15] compared the contents and features of different diabetic self-management mHealth applications. The study compares diabetes management applications in either iOS or Android	Content and features evaluation methods: • The search for mobile medical applications for diabetes is done in both the Apple iOS by using iPhone 6 and Google Android OS by using Lenovo A369i • It includes SMBG (self-monitoring of blood glucose) which helps the user in recording the measurement of blood glucose using the applications • The mobile applications are searched using a search key "diabetes"	• 35 applications were evaluated in both Apple and Google Play Store • Most applications were having features enabling the users for exporting the data files through email. Patients can store blood glucose readings and can later export them for review during their medical follow-up	According to the study, there are applications available for tracking a patient's medical records, particularly their blood glucose levels, as well as setting up reminders that can aid in improving diabetes control	Diabetes self-management is limited to only those patients who have the device and are technologically literate

(continued)

Table 1 (continued)

Objective of the paper	Methodology used	Result analysis	Advantages	Limitations
In this paper [16], usability evaluation of mHealth diabetes applications is performed for diabetic patients by utilizing a reformed heuristic evaluation method that includes (1) specialists of dual-domain (2) validated scenarios and user activity thorough assessments and ratings of severity factor	Heuristic evaluation method: The methodology involved: a. Expert evaluators: Experts were chosen based on their expertise with usability, their status as healthcare professionals, their knowledge of patient groups and their job requirements, and their knowledge of diabetes self-management b. User scenarios and tasks: The goal was to simulate how patients would use the system at a clinic or at home as part of a self-management approach. The panel was created which included physicians, registered nurses, and public health professionals c. Severity rating factors and scale: The heuristic evaluation (HE) methodology entails awarding a single severity score that is separated into frequency, impact, and persistence aspects	• There were 129 usability issues and 274 heuristic violations. Application views include Dashboard, Glucose Diary, Blood Pressure, and Medical Adherence View. Across these views, the number of usability issues ranges from 12 (low) to 34 (high) • Informaticists can use the modified HE to perform quick and resource-efficient heuristic processes. HE is employed in usability evaluation and is well suited to mHealth applications for chronic disease patients	• Dual-domain expertise brings some new and useful elements for evaluating usability • Dual-domain experts help improve user experience design • Dual-domain expertise can also assist in discovering patients' safety concerns that they may not be aware of	• Heuristics is complex • Highly skilled professionals are required for such evaluations
In this paper [17], the authors evaluated the usability of mHealth applications for health characteristics of public and health literate design methodology. The application cost was used to compare the health literacy aspects of diabetic apps (free or not) their languages, interfaces available, and user score	Expert-based usability evaluation methods: a. Diabetes-related keywords were typed into the Apple app store's search bar for iOS devices b. The information provided includes title, price, classification or category, rating related to age, total ratings in number given by application users, and star rating that range from 1 to 5 (with 5 being the best rating) c. Categories were created for three public health factors and four health literacy design methodologies or strategies. Types of diabetes, its continuum (diabetes behavior), and application emphasis were all categorized as public health factors. Type 1, Type 2, gestational diabetes, pre-diabetes, and others were several types of diabetes	• The paid applications were more useful in comparison to the free applications as they included more features • Paid apps used common everyday words. Paid apps also avoided the use of technical and medical terms that are not defined • Links were labeled more clearly in paid apps • Paid apps included a "back button." No other differences were there between paid and free apps related to organizing strategies	• The sample of diabetes-related applications coded was highly rated by the users and was appropriate for both children and adults • Resources are available to help improve the health literacy of the available mobile diabetes applications	• Diabetic applications which are paid were more likely to use common phrases used every day, and only premium or paid apps avoided using undefined technical and medical terminology. Paid applications displayed the content more clearly

(continued)

Table 1 (continued)

Objective of the paper	Methodology used	Result analysis	Advantages	Limitations
The authors of this [18] study conducted a systematic analysis of the diabetes applications which are available in iOS and Android and evaluated their functionality	Expert-based usability evaluation method: a. A search and screening technique for iOS and Android apps was developed with the help of information collected while using Apple and Google Play Store b. Categories "Health and Fitness" and "Medicine" were chosen for performing the search and screening strategy c. An expert-based usability assessment on the basis of a 10% sample of existing diabetes applications was performed as of April 2013 The chosen criteria were divided into main and sub-criteria, and experts used a Likert scale having 5 points to grade each sub-criterion. The mean of the sub-criteria was used to determine the key criteria ratings or scores	• In total, 656 applications were analyzed, of which 355 applications offered were only one function and 348 provided documentation function • On average, users gave it a 3.6-star rating. In the usability evaluation, 66 applications were evaluated • The "Comprehensibility" criterion was used to rate applications. The ability to read screen material aloud was available in 48 of the 66 applications	• Diabetes applications for age 50 and above were moderate to good in usability terms. The reviews concentrated on iOS and Android which are the most widely used operating systems	• The performance of applications having multifunctional capabilities was not satisfactory in terms of usability • This study did not take into account diabetes applications for Windows Phone or Blackberry • Paid applications were more organized
This paper [19] focuses on health information system assessment or evaluation methodologies that have developed from cognitive methods and usability engineering techniques	Heuristic evaluation and cognitive walkthrough method: usability testing and inspection are used Usability testing involves i. Exploratory tests carried out during the SDLC phase during preliminary design ii. Prototype testing carried out during requirement gathering iii. Validation checks to ensure that the software developed is suitable. Some inspection methods can be: • Heuristic assessment or evaluation—whether the user interface adheres to a collection of heuristics • Pluralistic walkthrough—a meeting in which researchers and users address usability issues • Cognitive walkthrough—stimulates the user's behavior and cognitive processes when performing tasks	Through the questionnaire, we were able to collect information such as the demographics, the users' age, and whether they use computers regularly A description of the task hierarchy is generated by cognitive task analysis. Video recording and screen capturing software, like Lotus Screen Cam, was used to capture the computer screens	• A variety of approaches to evaluate clinical information systems have also been developed and refined • Heuristic evaluation and cognitive walkthrough, the most popular evaluation methodologies, are used to assess healthcare systems • Advancements have been made for evaluating simulated or naturalistic events	• Some of the evaluation methods are time-consuming • It is challenging to integrate data acquired from different evaluation methodologies. Moreover, the potential relationship among the methods is not developed

(continued)

Table 1 (continued)

Objective of the paper	Methodology used	Result analysis	Advantages	Limitations
The authors of this study [20] evaluate the work that has been done in the field of user testing, with the aim of clarifying or specifying test procedures and identifying and improving resources to aid in the conduct of user tests. In this paper, user-based evaluation and Remote usability evaluation are performed	User-based evaluation and remote usability evaluation method: a. User-based evaluation is done involving the users to participate directly b. User performs tasks with the product and explores the product freely. Their behaviors are observed and recorded c. During the observation process, the time it takes to complete the assignment, the pace or rate at which it is done, and the types and number of errors are all captured and recorded d. Remote usability evaluation is performed in which the test participants are not present in the same location	• The study of the diaries takes into account the activities occurring in real-time environments. Participants record the activities • Computer screens are shared using video conferencing and sharing tools related to remote applications • The behaviors of the users are captured, collected, and then visualized as to which Web pages the people have explored	• It allows evaluators to interpret and visualize data • The events that are recorded can be transferred, i.e., exported to statistical tools for analysis • Since audio and video recording of user test sessions are done, it helps in indicating the area where the user faced problems • Both the user test and the empirical usability test are well documented	• Usability evaluations are costly • It is difficult to automate a process that eliminates the need for manual recording or logging • Tools for user test that includes data capture analysis of data and data representation are partially complete

(continued)

Table 1 (continued)

Objective of the paper	Methodology used	Result analysis	Advantages	Limitations
The study's [21] objective is to observe if an approach involving multi-method for collecting data and interpretation for patients' experiences with the mHealth applications related to self-management of Type 2 diabetes is feasible	Heuristic evaluation, cognitive walkthrough, and think-aloud methods: • Usability is assessed using heuristic evaluation and cognitive walkthrough which are the expert methodologies • Think-aloud method for assessing usability is used. Questionnaires and in-depth interviews are also used • The qualitative data is examined using Framework Analysis (FA). It assists in the study of descriptive, structured, and textual data sources, yielding reliable and qualitative results. There are five levels of FA 1. Familiarize the results with the details/data 2. Decide the framework or structure that is to be used 3. Create an index and use the structure to organize the results 4. Using a graph to display data 5. Mapping and interpreting the data	• Multiple kinds of data gathering methods result in a diverse set of issues related to usability and aid in the triangulation of data. Think-aloud usability testing was important for materializing the usability issues • The use of an in-depth interview and a questionnaire allows for data triangulation in the case of serious usability concerns • Data analysis using FA is a defined • UPT helped in an in-depth scheme classification and in the determination of severity rating score	• Methods of data gathering resulted in a thorough and full set of problems or issues related to usability, as well as the ability to triangulate (validate) the data • The structured data analysis helps in data validation to determine the most severe problems for the users • FA provides a feasible way to solve the usability problems from huge qualitative data • UPT aids to assess severity score or for the usability issue	• The study uses a randomized sample of diabetic patients drawn from a large and convenient sample. The frame of convenience sampling related to a big study makes it difficult to apply the observations to diabetic patients

(continued)

Table 1 (continued)

Objective of the paper	Methodology used	Result analysis	Advantages	Limitations
Authors in this paper [22], complete the evidence-based approach for assessing the mHealth system. The study serves as a model for well-organized usability techniques	Heuristic evaluation method: • The research used various methods recommended in ISO standards for assessing effectiveness, satisfaction, and efficiency including the SUS (System Usability Scale) instrument • Task completion success is used to determine effectiveness, followed by a count of the number of errors made during communication • The amount of effort and resources used by the users are also considered when determining performance • Satisfaction can be measured by the SUS instrument • The reliability, sensitivity, and validity of SUS are assessed • The SUS score is a number that varies from 0 to 100. A score of 70 or higher is considered suitable or fine, while a score of 35 or higher indicates a high usability degree level or an excellent score. A score of 50 or lower implies usability that is unacceptable	• The tasks related to export and fix values had the highest number of errors and the most amount of time spent on them • The average SUS satisfaction score of the system was 80.5, which signifies decent usability • According to the data, males finished the task more successfully and the young participants had high-performance scores • Microsoft Excel Spreadsheet was used. SPSS version 22 constructed and calculated descriptive statistics on effectiveness, efficiency, and satisfaction	• The analysis yielded a more comprehensive quantitative evaluation of usability • The findings provided information to the developers about the areas in which patients have problems and in which activities can be made perfect to suit or adjust users • The research is being used as an example for evaluating mHealth systems and wants to provide a benchmark to have repeatable and comparable outcomes	• This research used a small sample of users from a specific diabetes group, making it impossible to generalize the results to the general population • Since the participants are selected randomly, there is a fair chance of including the participants who are less or not literate
The authors of this paper [23], assess and compare the usability of three mobile applications for diabetes self-care	Think-aloud method: In this, usability testing is adopted for evaluating the diabetes application's usability. A total of 30 people with Type 2 diabetes (15 men and 15 women) took part in the usability evaluation. After completing task scenarios collection, participants tested the program with the aid of the scale known as System Usability Scale (SUS). Participants followed the think-aloud protocol while operating the applications	• The mean SUS of Application 3 was higher than the scores of Applications 1 and 2 • The high SUS and screenshots found throughout the think-aloud protocol process help in developing diabetes self-care applications	• The SUS score of Application 3 was higher than the scores of Application 1 and Application 2 • The advantages identified during the think-aloud protocol aid in the design of diabetes self-care	• Designers should take into account the concerns raised by elderly participants • The findings could be used to boost the usability of diabetes self-care apps

3 Objectives of the Study

The objectives of this paper are as follows:

- Identification of usability attributes apart from traditional Nielson [11] usability attributes, which can contribute to better usability assessment of mHealth applications.
- Comparing various usability evaluation methodologies adopted by researchers as shown in Table 1.
- Identifying the design flaws of mHealth applications and assisting them with guidelines in enhancing design solutions.

4 Usability Attributes

For assessing the mHealth applications, we commonly use the Nielson Model [11]. The attributes are discussed as follows.

- **Learnability:** When a user encounters a design for the first time, their learning ability can be characterized as how quickly and competently they can handle the basic activities. It also involves how efficiently the user can keep in mind the knowledge while usage.
- **Efficiency:** Efficiency measures how quickly and precisely the user can finish the task after being familiar with the design.
- **Memorability:** Memorability refers to how quickly users may regain expertise after being away from the design for a period of time.
- **Errors:** This refers to the frequency of errors produced, the seriousness or severity of the faults, and the methods for recovering.
- **Satisfaction:** The level of satisfaction with mHealth applications is measured by the satisfaction attribute. It refers to the comfort, likability, and pleasure of the user.

Apart from Nielson's attributes, in this work, we have identified some more attributes which contribute greatly to the significant usability evaluation of mHealth applications. These attributes are mentioned below:

- **Aesthetic design:** Aesthetic refers to the addition of simple and attractive features to the mHealth applications. The mHealth application should be designed so that it has pleasing qualities as well as simple to be used by all groups of users. The functionality of the mHealth applications can be enhanced by the inclusion of simple layouts which will fascinate the user.
- **Navigation**: Navigation of the different features of the mHealth applications should be very easy; otherwise, the users will find it difficult to evaluate the mHealth applications. If navigation is proper, it ensures that the user can navigate the mHealth applications without any problem.

- **Readability**: The content of the mHealth applications must be readable. Readability involves legibility and understandability. The mHealth applications should be legible in terms of color combinations, word style (italic, bold, etc.), font size, and typeface.
- **Cognitive Load:** The number of working memory resources (such as thinking, reasoning, and remembering) necessary to operate mHealth apps is referred to as cognitive load. The cognitive load of the mHealth applications should be reduced. Some ways in which cognitive load can be minimized are by removing the non-essential contents, breaking the content into smaller segments, presenting information both visually and verbally, etc.
- **Provision for Physically challenged users:** A group of users accessing mHealth applications can have certain physical disabilities such as hearing disability, motor disabilities, and vision problems. The user interface of the application should be able to manage these types of user groups as well.

5 Results and Analysis

For the literature review, the screening of different papers related to diabetes mHealth applications was carried out and eligible papers of this study were selected. Figure 1 depicts the flow diagram for the selection process of the included studies. Objectives, methodology used, result analysis, advantages, and limitations of 11 papers were discussed in this study. Figure 2 shows the graph of the papers reviewed in this study based on the years of their publications. By the year of publication, 9.1% of the studies ($n = 1$) were published between 2001 and 2005; 9.1% of studies ($n = 1$) between the year 2006–2010; 18.2% of the studies ($n = 2$) between the year 2011–2015; and 63.6% of the studies ($n = 7$) between the year 2016 and 2020. Through this study, a few more attributes are identified that include aesthetic design, navigation, readability, cognitive load, and provision for physically challenged users.

6 Conclusion and Future Work

For significant evaluation of usability, different attributes have been suggested by researchers which helps greatly for better performance and efficiency. Usability has been disintegrated into sub-attributes which are the hypothetical concepts for defining the success of the system. The involvement of the user also plays an important role to determine the usability of software applications once it has been made. A comparison study has been incorporated in this work which identifies the various attributes that play a significant role in evaluating mHealth applications and also describes the usability evaluation models used along with their advantages and limitations. In this paper, different UEMs were presented which helped in the identification of the most

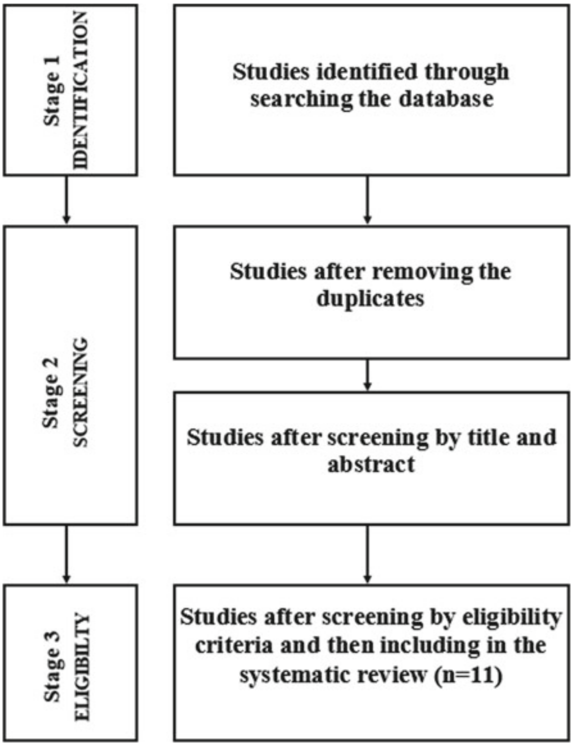


Fig. 1 Flow diagram for the selection process of included studies

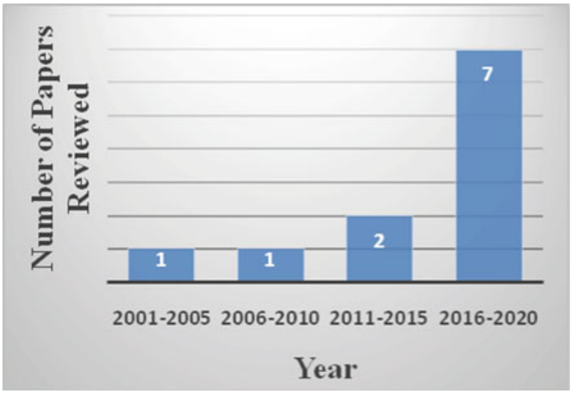


Fig. 2 Papers reviewed based on the years

important features relevant to the most suitable method for evaluating mHealth applications. Therefore, this paper will provide a pathway for the evaluators in choosing the most suitable UEM that may assist in evaluating a particular situation. This paper will be advantageous for the researchers and the students who are contributing in the software engineering field.

Future research to assure patients' adapting the mHealth applications should include the training needs of the patients, the problems they face while using the mHealth applications, and expedite its use for the long term. Inputs from the caregivers should be considered as they play a vital role in helping the patients to involve themselves in positive health behaviors while performing self-management.

References

1. Patrick K, Griswold WG, Raab F, Intille SS (2008) Health and the mobile phone. Am J Prev Med 35(2):177–181. https://doi.org/10.1016/j.amepre.2008.05.001
2. Ritterband LM, Andersson G, Christensen HM, Carlbring P, Cuijpers P (2006) Directions for the international society for research on internet interventions (ISRII). J Med Int Res https://doi.org/10.2196/jmir.8.3.e23
3. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM (2004) The effectiveness of web-based versus non-web-based interventions: a meta-analysis of behavioral change outcomes. J Med Int Res. https://doi.org/10.2196/jmir.6.4.e40
4. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M (2006) Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Int Res. https://doi.org/10.2196/jmir.8.2.e10
5. Pfaeffli L, Maddison R, Whittaker R, Stewart R, Kerr A, Jiang Y, Dalleck L (2012) A mHealth cardiac rehabilitation exercise intervention: findings from content development studies. BMC Cardiovasc Disord 12(1):1–9. https://doi.org/10.1186/1471-2261-12-36
6. American Diabetes Association. Standards of medical care in diabetes-2014. Diabetes Care, 37(SUPPL.1):14–80. https://doi.org/10.2337/dc14-S014
7. Maniam A, Dhillon JS (2015) Barriers to the effective use of diabetes self-management applications. The 3rd national graduate conference (NatGrad2015), April, pp 315–320
8. Mendiola MF, Kalnicki M, Lindenauer S (2015) Valuable features in mobile health apps for patients and consumers: content analysis of apps and user ratings. JMIR Mhealth Uhealth 3(2):e40. https://doi.org/10.2196/mhealth.4283
9. Harris MI (2000) Health care and health status and outcomes for patients with type 2 diabetes. Diabetes Care 23(6):754–758. https://doi.org/10.2337/diacare.23.6.754
10. Kahn JG, Yang JS, Kahn JS (2010) "Mobile" health needs and opportunities in developing countries. Health Aff 29(2):252–258. https://doi.org/10.1377/hlthaff.2009.0965
11. Nielson J, Mack R (1994) Usability inspection methods. Wiley, New York, NY
12. Fu HNC, Rizvi RF, Wyman JF, Adam TJ (2020) Usability evaluation of four top-rated commercially available diabetes apps for adults with type 2 diabetes. CIN: Comput, Inf, Nurs 38(6):274–280. https://doi.org/10.1097/cin.0000000000000596
13. Fontecha J, González I, Bravo J (2019) A usability study of a mHealth system for diabetes self-management based on framework analysis and usability problem taxonomy methods. J Ambient Intell Humanized Comput 1–11. https://doi.org/10.1007/s12652-019-01369-0
14. Jake-Schoffman DE, Silfee VJ, Waring ME, Boudreaux ED, Sadasivam RS, Mullen SP, Carey JL, Hayes RB, Ding EY, Bennett GG, Pagoto SL (2017) Methods for evaluating the content, usability, and efficacy of commercial mobile health apps. JMIR MHealth and UHealth 5(12). https://doi.org/10.2196/mhealth.8758
15. Izahar S, Lean QY, Hameed MA, Murugiah MK, Patel RP, Al-Worafi YM, Wong TW, Ming LC (2017) Content analysis of mobile health applications on diabetes mellitus. Front Endocrinol 8. https://doi.org/10.3389/fendo.2017.00318
16. Georgsson M, Staggers N, Weir C (2016) A modified user-oriented heuristic evaluation of a mobile health system for diabetes self-management support. CIN: Comput, Inf, Nurs, 34(2):77–84. https://doi.org/10.1097/cin.0000000000000209

17. Caburnay CA, Graff K, Harris JK, McQueen A, Smith M, Fairchild M, Kreuter MW (2015) Evaluating diabetes mobile applications for health literate designs and functionality, 2014. Preventing Chronic Dis 12. https://doi.org/10.5888/pcd12.140433
18. Arnhold M, Quade M, Kirch W (2014) Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 16(4):1–25. https://doi.org/10.2196/jmir.2968
19. Kushniruk AW, Patel VL (2004) Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 37(1):56–76. https://doi.org/10.1016/j.jbi.2004.01.003
20. Bastien JMC (2010) Usability testing: a review of some methodological and technical aspects of the method. Int J Med Informatics 79(4):e18–e23. https://doi.org/10.1016/j.ijmedinf.2008.12.004
21. Georgsson M, Staggers N (2016) An evaluation of patients' experienced usability of a diabetes mHealth system using a multi-method approach. J Biomed Inform 59:115–129. https://doi.org/10.1016/j.jbi.2015.11.008
22. Georgsson M, Staggers N (2015) Quantifying usability: an evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics. J Am Med Inform Assoc 23(1):5–11. https://doi.org/10.1093/jamia/ocv099
23. Hsieh, M.-H., Chen, Y.-C., & Ho, C.-H. (2019). A usability evaluation of diabetes mobile applications. Human-computer interaction. Design practice in contemporary societies, pp 3–15.https://doi.org/10.1007/978-3-030-22636-7_1

An Effective Diagnostic Framework for COVID-19 Using an Integrated Approach

Parul Agarwal, Sheikh Mohammad Idrees, Ahmed J. Obaid, Azmi Shawkat Abdulbaqi, and Sawsan Dheyaa Mahmood

Abstract The coronavirus, one of the deadliest virus erupted in Wuhan, China in December and has claimed millions of lives worldwide and infected too. This virus has off-late demonstrated mutations thus making it difficult for the health professionals to adopt a uniform means of cure. Many people due to lack of support have confined themselves at home. The hospitals too are running short of equipment and support systems. Thus, computational connectivity between the patients at home and the hospitals needs to be established. The objective of this paper is to propose a framework/model that connects all the stakeholders so that either in regular monitoring or in emergency cases help can be provided to them. It has been well established through research and case studies that critical factors associated with this disease are oxygen level (SPO2), pulse rate, fever, chest infection, cough causing choking, and breathlessness. Data shall be collected, stored, and analyzed for the above symptoms and for this cloud storage and blockchain technology would be used. It has been established through various studies that non-clinical techniques like AI and machine learning prove to be effective for the prediction and diagnosis of COVID-19. Using this theory as the standard basis, machine learning models like SVM, Naïve Bayes, and decision trees can be used for the analysis, diagnosis, and prediction. Using IoT and its variants, remote monitoring of patient, and consultation can be provided

P. Agarwal (✉)
Jamia Hamdard, New Delhi, India
e-mail: pagarwal@jamiahamdard.ac.in

S. M. Idrees
Norwegian University of Science and Technology, Trondheim, Norway
e-mail: sheikh.m.idrees@ntnu.no

A. J. Obaid
Faculty of Computer Science and Mathematics, University of Kufa, Kufa, Iraq
e-mail: ahmedj.aljanaby@uokufa.edu.iq

A. S. Abdulbaqi · S. D. Mahmood
College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq
e-mail: azmi_msc@uoanbar.edu.iq

S. D. Mahmood
e-mail: sawsan.d.mahmood@tu.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_11

to the patient. Appropriate action would be taken. In addition, a mobile application would enable the patients to gather or read about experiences of other patients. Thus, it would be established through the proposed framework, that an integrated approach of technologies has a great potential in such applications and offers several advantages.

Keywords Coronavirus · COVID-19 · Artificial intelligence · Blockchain · IoT · IoMT · CIoMT · Disease diagnosis · Treatment · Prediction · Mobile application · Smart health care

1 Introduction

Virus, a microscopic agent, are found in water, air, and soil. Common cold, flu, cough are the common infectious diseases caused by the virus. They enter our body and replicate through the normal cells, thus destroying them and make us sick. It largely depends on the immune system of our body, and the ability to fight with the virus or cause damage to us. When this virus spreads across at a large scale, worldwide disrupts the normal functioning of lives, which causes social and economic imbalance, and kills or infects lakhs, then the virus is termed as "pandemic virus". But for any viral infection, it is understood and medically proven that virus cannot be treated through antibiotics. Rather antibiotics work only for fungal or bacterial infections. This virus has found to be spreading more for some centuries now [1]. And the one that caused mass deaths was Spanish flu [2].

Coronavirus is another pandemic virus that was first reported in Wuhan, China in December, 2019 [3]. But, sadly was declared a pandemic only in [4]. Common symptoms of COVID-19 are similar to common flu which include fever, dry cough, sore throat, tiredness, body ache, and breathlessness (resulting in chest pain) [5].

Infected persons range from asymptomatic to those who show mild symptoms (say low-grade fever, mild cough, and tiredness) to those who show extreme symptoms which include drop in oxygen levels at several points of time and chest pain. The first category of asymptomatic patients are generally quarantined at home or at govt. run health-centers in absence of caregivers. The mild symptoms patient too can be remotely monitored at home or rushed to hospitals only in case of emergency if the parameters vary drastically. The third category is generally kept in hospitals to reduce the mortality rate and provide a life-saving support to them through physical monitoring by the health-care professionals. But, if the patient already suffers from acute or chronic illness like diabetes, respiratory diseases, or any type of respiratory or cardio-vascular disease then utmost care and regular attention and monitoring becomes a matter of concern. The severity and care may vary depending on the age [6] of the patient. Research works related to COVID-19 have been contributed and throw more light on various aspects related to pandemics which include pandemic prediction [7], drugs and medicines used [8–10], monitoring of the patient [11–13], and early detection and diagnosis [14–16]. This virus has indeed opened avenues for

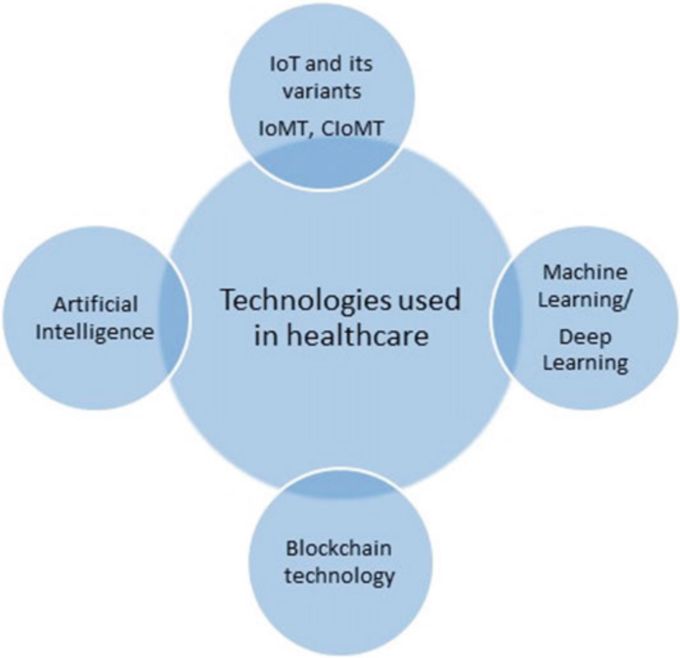


Fig. 1 Technologies related to health care

inter-disciplinary research where applications of science and technology in health care [17, 18] are much needed. These technologies as shown in Fig. 1 include big data analytics, machine learning, deep Learning, artificial intelligence (AI), IoT (Internet of Things), and its variants like IoMT (Internet of Medical Things), CIoMT (Cognitive Internet of Medical Things), and blockchain technologies. Some other works related to technologies and the area of focus in health care are mentioned in Table 1.

Currently, we all identify the best utilities of each of the technologies: IoT, blockchain, machine learning, and others. But no integrated framework exists that establishes meaningful relationships between each of them and can be used for translating the relationships into a form that would benefit the stakeholders of the health-care sector. IoT and its variants for data collection, blockchain for effective and secure data management, and machine learning for disease diagnosis by relying mainly on data and medical imaging, and Tele-health/ E-health for monitoring are effective techniques. This paper shall suggest a framework/ architecture that captures the essence of each of these techniques with a holistic view. This paper performs extensive research related to technologies related to health care, particularly, for COVID-19 in Sect. 2. Section 3 contains the proposed framework which adopts an integrated approach of several technologies so as to improve the health-care services, and its advantages are then discussed in Sect. 4. A conclusion of the paper is presented in Sect. 5.

Table 1 Research work related to technologies used and the area of focus

Reference No.	Area of focus	Technology used
[19]	Prediction	BP neural networks
[20]		ARIMA model
[21]		LSTM networks
[22]	Classification of confirmed cases and others	Artificial intelligence and regression analysis
[23]	Patient monitoring	Deep learning
[24]	Diagnosis and treatment	CT images-based deep learning approach
[25]		Blood tests-based machine learning approach
[26]	Remote health monitoring	A get-well loop program developed
[27]	Disease tracking	Data collection from mobile phones
[28]		Tracking app. Also includes a list of apps and some salient features of these apps for several countries around the globe
[29]	Drug discovery	Science and technology
[30]	Can be used for both treatment and remote monitoring	Sensor-based technology
[31]	Disease treatment	Blockchain technology
[32]	Reducing the workload of health-care workers	Artificial intelligence
[33]		Holistic view

2 Literature Review

Lot of research has been carried out related to COVID-19 and in particular on how to obtain the maximum benefits of using computing techniques for its prevention, prediction, cure, and remote monitoring. This section explores some of the related research work.

In this paper [34], the authors have examined the performance of deep learning algorithms with that of clinicians using medical imaging. Medical images have now been used for deep learning research. Particularly, convolutional neural network (CNN) take raw data as input, this network learns and develops its own representations for pattern recognition. Minimizing human interventions for feature selection, this network uses its own knowledge for feature selection depending on its importance and relevance. The measure of adherence to reporting standards, CONSORT for randomizes trials, and TRIPOD for non-randomized was used. Assessment of risk bias was done using Cochrane for randomized study and PROBAST tool for non-randomized study. Results related to these tools has been discussed in this work.

This paper, [35] used various classification models for identification of infected COVID-19 patients from the non-infectious ones to curb the spread of the disease. The study has been conducted using the X-ray images as the source data. The deep features are used for the classification. The data repositories include GitHub, Open-I, and Kaggle. A comparison of ResNet 50 and SVM with other classification models establishes this model better in terms of validation measures like F1 score, MCC, FPR, and Kappa.

Apostolopoulos and Mpesiana [36] proposes deep transfer learning (deep learning based) technique, for the prediction of infected COVID-19 patients automatically. 50 chest X-ray images of infected persons from GitHub repository and 50 chest X-ray images of healthy persons from Kaggle were used. As compared to three models used, it was shown through their study that ResNet50 achieved higher accuracy and can be used for early prediction for better performance. IoT and its applications in the health-care sector have been discussed in [37] this paper. Issues, solutions, and opportunities related to IoT and Health-care sector are discussed extensively. The authors also emphasize that IoT can be useful to governments in identifying solutions to problems that health-care sector faces and on a larger scale can benefit the society. Several works related to IoT and Health-care are discussed in [38–43]. AI has been used extensively in several sectors including the health care. In particular, its contribution in dealing with the pandemic COVID-19 has been extraordinary. Researchers are harnessing its benefits for prediction, detection, classification, cure, and diagnosis. 13 groups of problems related to the pandemic are identified and AI can be used for finding solutions to these problems. Data sources related to corona are presented in [44] for future use and research. The authors discuss that use of blood tests, chest X-rays and CT scans [30] are the methods being adopted worldwide for the detection of the novel coronavirus. But the need is to provide low-cost solutions for the same. To enable this, they propose an AI-based framework which is uses built-in smartphone sensor-based technology for the disease detection. They emphasize that since smartphones are a day-to-day utility owned by all, so the sensors installed in smartphones would provide data which can be analyzed and processed for identifying the severity of the disease. An AI-based framework has been suggested in [45] for the identification of COVID-19 cases. The framework uses survey data gathered through smartphones when people cannot move out of the houses as cities are under quarantine.

In this paper, [46] authors extensively examine the trends in China and suggest a model, namely Susceptible-Exposed-Infectious-Removed (SEIR) for deriving epidemic curve. This helps to correlate the peak of cases and the fall with time. They have also suggested an AI approach, for the prediction of the epidemic. This study [47] proposes and generates compounds against SARS-COV-2, using a drug discovery pipeline which is deep learning based. Identification or screening of patients at an initial stage [48] is achieved in this study using a deep learning approach. CT images/scans are used as the basis for distinguishing coronavirus patients from Influenza-A pneumonia and these from the normal ones. Through this work, they have explored CT images as an important diagnostic tool for this identification and can be helpful for the clinical trials by the doctors. A deep learning-based approach called COVIDX-Net [49] has been proposed in this work. This approach uses X-ray images as the diagnostic tool which can help the radiologists in identification of COVID-19 patients. A few classifiers have been used for the experiments and for evaluation. The study [50] aims at early detection of coronavirus using machine learning techniques. Abdominal CT images were used for this study and analysis. Since, COVID-19 exhibited different behaviors from viral pneumonia, so detection of coronavirus becomes essential. Patches of different sizes from 150 images were taken. For classification, feature extraction was then applied on these patches. Four different algorithms were used for feature extraction. Assessment of results was done using several parameters which include precision, f-score, accuracy, sensitivity, and specificity. The best results were obtained with gray level size zone matrix algorithm. The next section discusses the framework that would be effective and improve the health care particularly for the COVID-19 infected/ suspected patients.

3 Proposed Framework

The proposed architecture encompasses the advanced technologies as discussed in Fig. 1 at various steps and suggests how integration of these techniques for various domain applications like diagnosis, storage, prediction, circulation, treatment, and cure can help us develop a model well-suited for all in the current situation. The steps of the model are mentioned below.

Step 1. Data collection: Wearable technologies can incorporate sensors which be used for capturing acoustic information (volume and frequency), biological information (temperature, heart-rate, fatigue, stress, and respiration rate), optical information (brightness and refraction), and other environmental information (temperature and humidity). Data using biological sensors various important symptoms which include ECG monitors, pulse oximeters, fatigue, cough, fever, and tiredness. Techniques for these are discussed in [51–55]. Other parameters like travel history/get-togethers or meetings attended along with the dates would be collected. Other data collected using input from the end-user would be personal details (age is an important parameter), and medical history with critical ones being diabetes, blood pressure, and heart disease.

Step 2. Data storage: Thus, we understand that there are broadly two types of data: personal data and electronic health records. All the data collected (known as Data Lake), is then encrypted, digitally signed and then stored as a health blockchain. Each digital record also stores the location of the health data it references. It also contains the patient's public ID. The patient's private key links their identity to blockchain data. The private key can be shared with health organizations if the person desires by giving desired access permissions. Using this key, the data can be decrypted to uncover the original patient's data. Blockchain offers advantages (most relevant being security, privacy, and accessibility) that can revolutionize the way health-care sector works, so it can be harnessed for data storage and data sharing [56] as well.

Step 3. Data preprocessing: As discussed, in Step 1, the data collected would be heterogenous in nature and may have several features which may be irrelevant. Thus, conversion of data to a standardized format, reducing the size by selecting interesting data [57] and removing of outliers are performed as part of this step.

Step 4. Disease diagnosis, treatment, and cure.

Each of the patient would fall any of the three cases:

- Suspected (through contact tracing would be quarantined in centers/or at homes).

  Through live streaming of CT scans and X-ray images further processed by AI-based sensors can help the lab technicians for further analysis. If they show symptoms in about five days, which is the average incubation period, then would become the confirmed positive and dealt accordingly.
- Confirmed positive (depending on age and other related ailments and depending on severity of disease would be either kept at home or sent to hospital). If kept at home then, a government physician can be assigned and local dispensaries can monitor the health. Remote health monitoring techniques using CIoMT can be used. Patient feedback and response can be used for effective treatment using CIoMT. Also, geographically driven real-time identification of positive cases can be done as the health-care units are connected via IoT. Mobile apps giving updated information area wise about houses or people reported as positive using IoT. Drugs/medicines delivery can be ensured by the blockchain companies. If critical parameters report an absurd rise or fall, then will be sent to hospital.
- Confirmed negative No major action needs to be taken. Just precautions need to be continued to prevent the disease from infecting this person. Not only prediction, the model can be used for identifying the treatment response which can then be utilized to treat other patients by self-learning and by self- adaptation.

Machine learning models like SVM, decision tables, and neural networks can be applied for the disease diagnosis. The reason for choosing machine learning models are plenty. Machine learning, an advanced concept of AI, gives strategic way for the development of complex, algorithmic methods and uses statistical techniques for bio-medical data analysis. Then hidden patterns or knowledge can be extracted from past experience. Performance evaluation of ML models is then followed by its optimization by using new rules. In [58], the authors used various ML techniques like decision trees, logistic regression, support vector machine, Naïve-Bayes for prediction of COVID-19 infection for negative and positive cases of an epidemiology dataset for patients based in Mexico. It was shown through experiments that with respect to accuracy, decision tree was the best among others. It also indicated that age was an important feature among other features. The model also indicated that those suffering from chronic diseases like diabetes, hypertension, asthma, etc., are likely to be infected. Also, males are more prone to the infection than women. Also, this study proved that decision trees help with the diagnosis of suspected COVID-19 patients. Another study [59–62] also proved that decision tree model is best suited for diagnosis and also for predicting the possibility of recovery of COVID infected patients. These

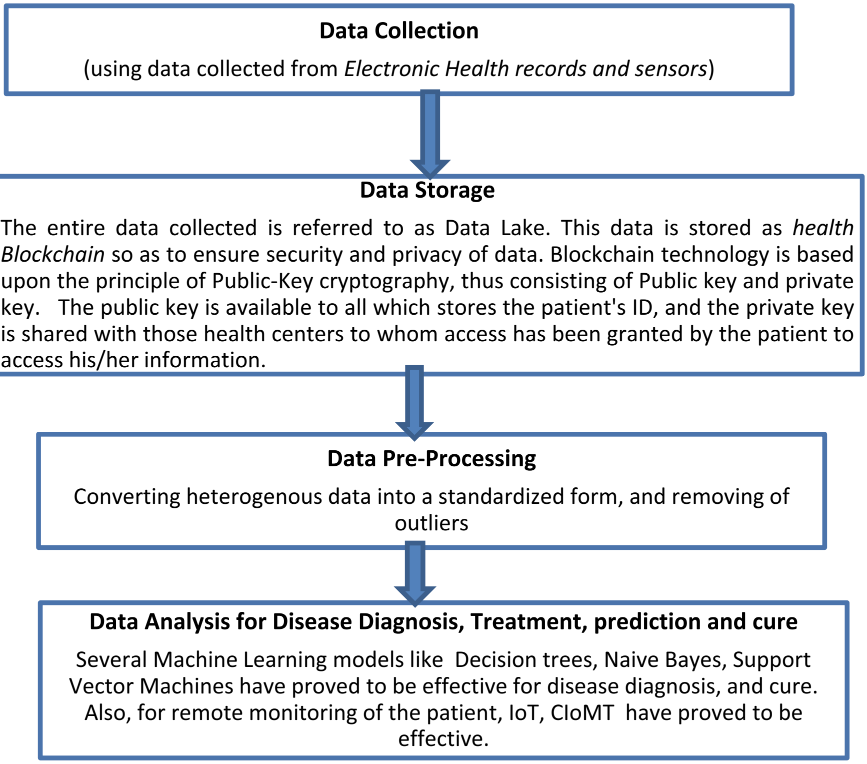


Fig. 2 Proposed framework

studies were used as the basis for choosing ML models for the diagnosis and recovery in the current study. Figure 2 depicts the proposed framework.

Other prominent features of the framework are:

- A restricted entry to the hospital/ quarantine centers can be ensured by making the entry points micro-chip enabled or eye retina-based entry would be granted to only authorized persons.
- The room would adjust the room temperature using a smart electrical system. This would not only provide comfort to the patient but also, reduce power consumption. An efficient energy management can be achieved using this.
- In case, a patient is old and incapable to talk to his family members or do video calls, then the nursing staff would mandatorily provide this service.
- The medical staff, especially in case of COVID-19 patients avoid visiting them to avoid infection. This would be monitored remotely by using AI-based approach which would ensure regular attendance monitoring of the staff.
- The clinical reports are not shared by the hospitals with the family members and only handed over to them at the time of discharge. As a result, they are unaware

of the real situation and health of the family member. Blockchain would ensure data sharing and enable the family members to monitor the health remotely and enable them to take second opinion from any other doctor if they want. This same process can be adopted for treatment and medicines being given to the patient. Also, sharing of readings of several important features like sugar level, blood pressure reading, pulse rate, and SPO2 can be shared with them. This would ensure transparency at all levels.

- A constant video monitoring of the patient's room should be ensured to swiftly identify deterioration in health and immediate support by the medical staff.
- The patients inside the hospital/ quarantine center would be given a wearable sensor. This sensor would continuously monitor critical parameters like SPO2, BP, sugar, and pulse rate. In case of any sudden rise or fall of these values, an emergency signal can be sent to the control room of the health center.
- Depending on the health condition of the patient and sensors which identify the taste habits of a patient, a personalized menu plan can be crafted and automatically updated with the approval of the doctor.

A mobile application that would connect people across the globe at one platform, and each one of them can share their journey of recovery with each other. This would not only help each other but also ensure a mental wellbeing using a digital solution [60]. It has been observed that not only, physical health monitoring of the patient, but COVID-19 can have a negative impact on the mental health too. Success stories of people around can motivate a person and bring positivity. These factors are as relevant and important in the recovery of a patient as are the physical symptoms.

4 Discussion

This framework does not work on any one technology related to health care, rather uses an integrated approach and derives the benefit of each one of them. In addition to the benefits that can be derived from the proposed framework, blockchain companies can be used for the timely delivery of medicine. Machine learning models (an efficient diagnostic model) can be effectively used for diagnosis and prediction. IoT, IoMT, and CIoMT, which create an inter-connected network of the stakeholders (like hospitals/government centers/nurses/lab technicians) can be used effectively to monitor the patient either remotely or otherwise. These technologies would trap the feedback and patient response for treatment. It would also help the doctors to swiftly and effectively respond to new cases encountered. The proposed framework would not only reduce the impact of the disease, the morality rate would be decreased and effective follow-ups can be done. A lot of research has been done related to COVID-19, yet, it lacks an in-depth and complete insight of several issues that stakeholders are facing in real world.

5 Conclusion

This paper has tried to visualize a framework which can be used for effective prediction, diagnosis, treatment, and cure. It will also help the physicians to effectively monitor new cases based on the learning capability of machine learning models. For future work, this framework can be implemented and can be used for handling COVID-19 in an effective manner. Advanced deep-learning models/techniques might prove to be more effective in future for tackling this disease. Effective drug discovery and vaccines, are the need of the hour. It has also been mentioned in the paper that there is a need of technological innovations that monitor the mental wellbeing of the patients. These are the challenges that open avenues for future research work.

References

1. Tumpey T, García-Sastre A, Taubenberger J, Palese P, Swayne D, Pantin-Jackwood M, Schultz-Cherry S, Solórzano A, Van Rooijen N, Katz J, Basler C (2005) Pathogenicity of influenza viruses with genes from the 1918 pandemic virus: functional roles of alveolar macrophages and neutrophils in limiting virus replication and mortality in mice. J Virol 79:14933–14944. https://doi.org/10.1128/JVI.79.23.14933-14944.2005
2. Erkoreka A (2010) The Spanish influenza pandemic in occidental Europe (1918–1920) and victim age. Influenza Other Respir Viruses 4:81–89
3. Xu Y, Li X, Zhu B, Liang H, Fang C, Gong Y, Guo Q, Sun X, Zhao D, Shen J, Zhang H, Liu H, Xia H, Tang J, Zhang K, Gong S (2020) Characteristics of pediatric SARS-CoV-2 infection and potential evidence for persistent fecal viral shedding. Nat Med 26(4):502–505. https://doi.org/10.1111/j.1750-2659.2009.00125.x
4. World Health Organization (2020) 'WHO Director-General's opening remarks at the media briefing on COVID-19–11 March 2020. https://www.who.int/dg/speeches/detail/whodirector-general-s-opening-remarks-at-the-media-briefingon-covid-19-11-march-2020
5. Rothan HA, Byrareddy SN (2020) The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. J Autoimmun 109:102433. https://doi.org/10.1016/j.jaut.2020.102433
6. Liu K, Chen Y, Lin R, Han K (2020) Clinical features of COVID-19 in elderly patients: a comparison with young and middle-aged patients. J infect 80(6):e14–e18
7. Davenport T, Kalakota R (2019) The potential for artificial intelligence in healthcare. Future Healthcare J 6(2):94–98. https://doi.org/10.7861/futurehosp.6-2-94
8. Sohrabi C, Alsafi Z, O'Neill N, Khan M, Kerwan A, Al-Jabir A, Iosifidis C, Agha R (2020) World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). Int J Surg 76:71–76. https://doi.org/10.1016/j.ijsu.2020.02.034
9. Chen DS, Yang J, Yang W, Wang C, Bärnighausen T (2020) COVID-19 control in China during mass population movements at New Year. Lancet 395(10226):764–766
10. Bobdey S, Ray S (2020) Going viral–COVID-19 impact assessment a perspective beyond clinical practice. J Mar Med Soc 22(1):9
11. Haleem A, Vaishya R, Javaid M, Khan IH (2020) Artificial Intelligence (AI) applications in orthopaedics: an innovative technology to embrace. J Clin Orthop Trauma 11(Suppl 1):S80–S81.https://doi.org/10.1016/j.jcot.2019.06.012
12. Biswas K, Sen P (2020) Space-time dependence of coronavirus (COVID-19) outbreak. arXiv preprint arXiv:2003.03149. Mar 6
13. Stebbing J, Phelan A, Griffin I, Tucker C, Oechsle O, Smith D, Richardson P (2020) COVID-19 Combining antiviral and anti-inflammatory treatments. The Lancet. Infect Dis 20(4):400–402.https://doi.org/10.1016/S1473-3099(20)30132-8

14. Hu Z, Ge Q, Jin L, Xiong M (2020) Artificial intelligence forecasting of COVID-19 in China. arXiv preprint arXiv:2002.07112. Feb 17
15. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, Tao Q, Sun Z, Xia L (2020) Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology 296(2):E32–E40. https://doi.org/10.1148/radiol.2020200642
16. Luo H, Tang QL, Shang YX, Liang SB, Yang M, Robinson N, Liu JP (2020) Can Chinese medicine be used for prevention of corona virus disease 2019 (COVID-19)? A review of historical classics, research evidence and current prevention programs. Chin J Integr Med 26(4):243–250. https://doi.org/10.1007/s11655-020-3192-6
17. Agrebi S, Larbi A (2020).Use of artificial intelligence in infectious diseases. A I Prec Health 415–438.https://doi.org/10.1016/B978-0-12-817133-2.00018-5
18. Tuli S, Tuli S, Wander G, Gill P, Singh D, Schahram S, Rizos R, Omer (2019) Next generation technologies for smart healthcare: challenges, vision, model, trends and future directions. Internet Technol Lett 3:e145. https://doi.org/10.1002/itl2.145
19. Bai Y, Jin Z (2005) Prediction of SARS epidemic by BP neural networks with online prediction strategy. Chaos, Solitons Fractals 26(2):559–569
20. Benvenuto D, Giovanetti M, Vassallo L, Angeletti S, Ciccozzi M (2020) Application of the ARIMA model on the COVID-2019 epidemic dataset. Data Brief 29:105340. https://doi.org/10.1016/j.dib.2020.105340
21. Chimmula VKR, Zhang L (2020) Time series forecasting of COVID-19 transmission in Canada using LSTM Networks. Chaos, Solitons Fractals. https://doi.org/10.1016/S0140-6736(20)30421-9
22. Pirouz B, Shaffiee HS, Piro P (2020) Investigating a serious challenge in the sustainable development process: analysis of confirmed cases of COVID-19 (new type of coronavirus) through a binary classification using artificial intelligence and regression analysis. Sustainability 12:2427
23. Gozes O, Frid AM, Greenspan H, Browning PD, Zhang H, Ji W, Bernheim A, Siegel E (2020) Rapid AI development cycle for the Coronavirus (COVID-19) pandemic: initial results for automated detection & patient monitoring using deep learning CT image analysis. arXiv preprint arXiv:2003.05037. Mar 10
24. Wang S, Kang B, Ma J et al (2020) A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). medRxiv. https://doi.org/10.1101/2020.02.14.20023028
25. Wu F, Wu T, Yuce MR (2018) An Internet-of-Things (IoT) network system for connected safety and health monitoring applications. Sensors 19(1):21. https://doi.org/10.3390/s19010021
26. Annis T, Pleasants S, Hultman G, Lindemann E, Thompson JA, Billecke S, Badlani S, Melton GB (2020) Rapid implementation of a COVID-19 remote patient monitoring program. J Am Med Inform Assoc 27(8):1326–1330. https://doi.org/10.1093/jamia/ocaa097
27. BBC: Coronavirus: Israel enables emergency spy powers. https://www.bbc.com/news/technology-51930681
28. MIT: Covid Tracing Tracker - a flood of coronavirus apps are tracking us. Now it's time to keep track of them. https://www.technologyreview.com/2020/05/07/1000961/launching-mittr-covid-tracing-tracker/
29. Ekins S, Mottin M, Ramos PRPS, Sousa BKP, Neves BJ, Foil DH, Zorn KM, Braga RC, Coffee M, Southan C, Puhl AC, Andrade CH (2020) Déjà vu: stimulating open drug discovery for SARS-CoV-2. Drug Discovery Today 25(5):928–941
30. Maghdid HS, Ghafoor KZ, Sadiq AS, Curran K, Rabie K (2020) A novel AI-enabled framework to diagnose coronavirus COVID 19 using smartphone embedded sensors: design study. arXiv preprint: 2003.07434
31. Kalla A, Hewa T, Mishra RA, Ylianttila M, Liyanage M (2020) The role of blockchain to fight against COVID-19. IEEE Eng Manage Rev 48(3). https://doi.org/10.1109/EMR.2020.3014052
32. Ting D, Carin L, Dzau V, Wong TY (2020) Digital technology and COVID-19. Nat Med 26(4):459–461. https://doi.org/10.1038/s41591-020-0824-5
33. Wan KH, Huang SS, Young AL, Lam D (2020) Precautionary measures needed for ophthalmologists during pandemic of the coronavirus disease 2019 (COVID-19). Acta Ophthalmol 98(3):221–222. https://doi.org/10.1111/aos.14438

34. Nagendran M, Chen Y, Lovejoy CA, Gordon A, C, Komorowski M, Harvey H, Topol EJ, Ioannidis JPA, Collins GS, Maruthappu M (2020) Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies in medical imaging. BMJ 368:1–12
35. Kumar P, Kumari S (2020) Detection of coronavirus disease (COVID-19) based on deep features. https://www.Preprints.Org/Manuscript/202003.0300/V1. March, p 9
36. Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 43(2):635–640. https://doi.org/10.1007/s13246-020-00865-4
37. Usak M, Kubiatko M, Shabbir MS, Dudnik OV, Jermsittiparsert K, Rajabion L (2020) Health care service delivery based on the internet of things: a systematic and comprehensive study. Int J Commun Syst 33(2):e4179. https://doi.org/10.1002/dac.4179
38. Wu J, Zhang P, Zhang L, Meng W, Li J, Tong C, Li Y, Cai J, Yang Z, JZhu J, Zhao M, Huang H, Xie X, Li S (2020) Rapid and accurate identification of COVID-19 infection through machine learning based on clinical available blood test results. https://doi.org/10.1101/2020.04.02.20051136
39. Hamidi H (2019) An approach to develop the smart health using internet of things and authentication based on biometric technology. Future Gener Comput Syst 91:434–449
40. Rath M, Pattanayak B (2019) Technological improvement in modern health care applications using Internet of Things (IoT) and proposal of novel health care approach. Int J Hum Rights Healthcare 12(2):148–162. https://doi.org/10.1108/IJHRH-01-2018-0007
41. Darwish A, Hassanien A, Elhoseny M, Sangaiah AK, Muhammad K (2019) The impact of the hybrid platform of internet of things and cloud computing on healthcare systems: opportunities, challenges, and open problems. J Ambient Intell Humaniz Comput 10:4151–4166
42. Zhong CL, Li YL (2020) Internet of things sensors assisted physical activity recognition and health monitoring of college students. Measurement 159
43. Din S, Paul A (2019) Smart health monitoring and management system: toward autonomous wearable sensing for Internet of Things using big data analytics. Futur Gener Comput Syst 91:11–619. https://doi.org/10.1016/j.future.2017.12.059
44. Thi NT (2020) Artificial Intelligence in the battle against coronavirus (COVID-19): a survey and future research directions. figshare. Preprint. https://doi.org/10.6084/m9.figshare.12127020.v6
45. Rao SA, Vazquez JA (2020) Identification of COVID-19 can be quicker through artificial intelligence framework using a mobile phone-based survey when cities and towns are under quarantine. Infect Control Hosp Epidemiol 41(7):826–830. https://doi.org/10.1017/ice.2020.61
46. Yang Z, Zeng Z, Wang K, Wong SS, Liang W, Zanin M, Liu P, Cao X, Gao Z, Mai Z, Liang J, Liu X, Li S, Li Y, Ye F, Guan W, Yang Y, Li F, Luo S, Xie Y, He J (2020) Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. J Thoracic Dis 12(3):165–174 https://doi.org/10.21037/jtd.2020.02.64
47. Zhavoronkov A, Aladinskiy V, Zhebrak A, Zagribelnyy B, Terentiev V, Bezrukov DS et al (2020) Potential COVID-2019 3C-like protease inhibitors designed using generative deep learning approaches. ChemRxiv. Preprint. https://doi.org/10.26434/chemrxiv.11829102.v2
48. Xiaowei X, Xiangao J, Chunlian M, Peng D, Xukun L, Shuangzhi L, Liang Y, Qin N, Yanfei C, Junwei S, Guanjing L, Yongtao L, Hong Z, Jun L, Kaijin X, Lingxiang R, Jifang S, Yunqing Q, Wei W, Tingbo L, Lanjuan L (2019) A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineeringhttps://doi.org/10.1016/j.eng.2020.04.010
49. Hemdan EED, Shouman MA, Karar ME (2020) COVIDX-Net: a framework of deep learning classifiers to diagnose COVID-19 in X-ray images. https://arxiv.org/abs/2003.11055
50. Barstugan M, Ozkaya U, Ozturk S (2020) Coronavirus (COVID-19) Classification using CT Images by machine learning methods 5:1–10. Arxiv: arXiv:2003.09424
51. Medina J, Espinilla M, Fernández ALG, Martínez L, Intelligent multi-dose medication controller for fever: from wearable devices to remote dispensers. Comput Electr Eng 65:400–412. https://doi.org/10.1016/j.compeleceng.2017.03.012

52. Umayahara Y, Soh Z, Sekikawa K, Kawae T, Otsuka A, Tsuji T (2018) A mobile cough strength evaluation device using cough sounds. Sensors 18(11):3810. https://doi.org/10.3390/s18113810
53. Ichwana D, Ikhlas RZ, Ekariani S (2018) Heart rate monitoring system during physical exercise for fatigue warning using non-invasive wearable sensor. In: 2018 International conference on information technology systems and innovation (ICTSI), Bandung - Padang, Indonesia, pp 497–502
54. Askarian B, Yoo SC, Chong JW (2019) Novel image processing method for detecting strep throat (streptococcal pharyngitis) using smartphone. Sensors 19(15):3307. https://doi.org/10.3390/s19153307
55. Gaidhani A, Moon KS, Ozturk Y, Lee SQ, Youm W (2017) Extraction and analysis of respiratory motion using wearable inertial sensor system during trunk motion. Sensors 17(12):2932. https://doi.org/10.3390/s17122932
56. Fan K, Wang S, Ren Y, Li H, Yang Y (2018) MedBlock: efficient and secure medical data sharing via blockchain. J Med Syst 42(8):136. https://doi.org/10.1007/s10916-018-0993-7
57. Li Y, Hai T, Z, Jeorge G et al (2020). A machine learning-based model for survival prediction in patients with severe COVID-19 infection medRxiv 2020.02.27.20028027. https://doi.org/10.1101/2020.02.27.20028027
58. Muhammad LJ, Algehyne EA, Usman SS et al (2021) Supervised machine learning models for prediction of COVID-19 infection using epidemiology dataset. SN Comput Sci 2:11. https://doi.org/10.1007/s42979-020-00394-7
59. Muhammad LJ, Islam MM, Usman SS et al (2020) Predictive data mining models for novel coronavirus (COVID-19) infected patients' recovery. SN Comput Sci 1:206. https://doi.org/10.1007/s42979-020-00216-w
60. Chew AMK, Ryan O, Hsien LH, Mallika R, Grisan KV, Verma SK, Fung DSS, Sheng S, Leong JJ, Gunasekeran DV (2020) Digital health solutions for mental health disorders during COVID-19. Front Psychiatry 11:898.https://doi.org/10.3389/fpsyt.2020.582007
61. Shahzad F, Abid F, Obaid A, Kumar Rai B, Ashraf M, Abdulbaqi A (2021) Forward stepwise logistic regression approach for determinants of hepatitis B & C among Hiv/Aids patients. Int J Nonlinear Anal Appl 12(Special Issue):1367–1396. https://doi.org/10.22075/ijnaa.2022.5717
62. Kautish S, Peng S-L, Obaid AJ (2021) Computational intelligence techniques for combating COVID-19. Springer International Publishing

Diabetes Mellitus Prediction Through Interactive Machine Learning Approaches

Rabinarayan Panda, Sachikanta Dash, Sasmita Padhy, and Rajendra Kumar Das

Abstract Diabetes is a long-term illness that has the potential to disrupt the global healthcare system. Based on the survey report of International Diabetes Federation (IDF), there are around 382 millions of people, who are affected by diabetes worldwide. This number will have increased to 592 million by 2035. Diabetes is a disease characterized by an increase in blood glucose levels. Elevated blood glucose is characterized by frequent urination, increased thirst and increased hunger. Diabetic consequences include kidney failure, blindness, heart failure, amputations and stroke, to name a few. When we ingest food, our bodies turn it into sugars or glucose. Machine learning is a new field of data science that investigates how computers learn from their prior experiences. The objective of this study is to develop a system that can detect diabetes in a patient early and more accurately using a combination of machine learning techniques. The objective of this study is to use four supervised machine learning algorithms to predict diabetes: Support Vector Machine, logistic regression, random forest and k-nearest neighbour. Each algorithm is used to calculate the model's accuracy. The model with the best accuracy for predicting diabetes is then picked. This paper proposes a comparative study for accurately predicting diabetes mellitus. This research also aims to develop a more efficient approach for identifying diabetic disease.

Keywords KNN · Logistic regression · Random forest · SVM · ROC

R. Panda
Department of CSE, GIET University, Gunupur, India

S. Dash (✉) · R. K. Das
Department of CSE, DRIEMS, Cuttack, India
e-mail: dash.sachikanta@gmail.com

S. Padhy
Department of CSE, KL University, Vaddeswaram, India

R. K. Das
Department of ENTC, DRIEMS, Cuttack, India

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_12

1 Introduction

This article requires certain precise actions related to diabetes control and prevention. Previously, statistics showed that one out of every 10 people in the USA had diabetes. Regardless, it is expected that by 2045, it will be able to assist one out of every three people. This is a real problem that has to be addressed. When the blood glucose concentration rises to dangerously high levels, diabetes becomes a chronic condition. This is a major cause of various problems and illnesses [1], such as renal disease and heart disease. Diabetic predisposition is also caused by a variety of bad food habits and a lack of proper bodily routines. The World Health Organization (WHO) has said that the total number of people living with diabetes has skyrocketed in recent years. Managing multiple diabetic data sets is required to enhance the present rate of diabetes patients and reduces it to an absolutely inconsequential level by focusing on reducing it on a large scale. Certain approval techniques are also incorporated to operate with the diabetes forecast project's pure precision.

Diabetes is a rapidly spreading disease that affects individuals of all ages, including children. To understand diabetes and how it develops, we must first understand what happens in the body without diabetes. Carbohydrate meals are our body's primary source of energy. Bread, cereal, pasta, rice, fruit, dairy products and vegetables are all carbohydrate foods (especially starchy vegetables). Our bodies transform these nutrients to glucose when we consume them. Glucose flows throughout the body in the bloodstream. Some of the glucose is delivered to our brain, where it helps us think and function correctly. The rest of the glucose is converted to energy that is transported and used by our body's cells. Also liver utilizes this energy for storing purposes that can be utilized latter. Insulin is required for the body to use glucose as an energy source. Insulin is a hormone produced by beta cells in the pancreas. Insulin works in a similar way as a key in a lock. Insulin attaches to cell doors and opens them, allowing glucose to flow from the circulation into the cell through the cell door. When the pancreas is unable to create enough insulin or the body is unable to use the insulin it does make (insulin resistance), glucose builds up in the bloodstream, causing hyperglycemia and diabetes. The presence of excessive quantities of sugar (glucose) in the blood and urine is a symptom of diabetes mellitus.

1.1 Type of Diabetes with Symptoms

There are three kinds of diabetes found in human body. These include Type 1, Type 2 and gestational diabetes. Type 1 diabetes occurs when the immune system is weakened, and the cells are unable to create enough insulin. Type 2 diabetes occurs when the cells generate insufficient insulin or when the body fails to utilize the insulin properly. Pregnant women who acquire high-blood sugar levels quickly develop gestational diabetes. The common symptoms found in diabetes patients include:

- Urination on a regular basis

- Increase in thirst
- Tired/Sleepiness
- Loss of weight
- Distorted eyesight
- Emotional ups and downs
- Perplexity and inability to concentrate
- Infections are common.

2 Literature Survey

In this healthcare area, a variety of computing approaches were utilized. The application of several machine learning algorithms for predicting diabetes mellitus is the subject of this literature review. We extract information from the provided medical data in order to achieve flawless accuracy. Md. Faisal Faruque et al. [2] used the random forest method to create a predictive analytic model. Chaki et al. [3] utilized tenfold cross-validation as an assessment technique for three distinct algorithms: decision tree, Naive Bayes and SVM, with Naive Bayes outperforming the other algorithms by 75%.

To forecast diabetes mellitus in its early stages, Alam et al. [4] utilized random forest, KNN, Naive Bayes, SVM and decision tree. We are now using machine learning algorithms and statistical data in the healthcare area to comprehend the sick data that has been discovered. Because the machine learning area encompasses a wide range of approaches and studies, it is difficult to establish a comparison based on which algorithm is faster at producing prediction results. The algorithm's categorization was not tested using the cross-validation approach. Different data mining approaches were utilized to predict and assess diabetic mellitus. We used real-world data sets by gathering information from the supplied data sets since we employed three data mining approaches.

VijiyaKumar et al. [5] presented the random forest method using machine learning techniques. The suggested model produces the best diabetic prediction results, demonstrating quickly predicting diabetes mellitus.

Predicting diabetes onset: an ensemble supervised learning technique was reported by Nnamoko et al. [6]. For the ensembles, five commonly used classifiers are utilized, and their outputs are aggregated using a meta-classifier. The findings are reported and compared to other research in the literature that used the same data set. It is demonstrated that diabetes onset prediction can be done more accurately utilizing the proposed technique.

Diabetes prediction using machine learning techniques, given by Joshi et al. [7], seeks to predict diabetes using three distinct supervised machine learning methods: SVM, logistic regression and ANN. This study provides an excellent method for detecting diabetes illness sooner. Abdulqadir et al. [8] proposed employing data mining to build an intelligent diabetic disease prediction system that provides analysis of diabetes malady using a database of diabetes patients. In this approach, they suggest using Bayesian and KNN (k-nearest Neighbour) algorithms to a diabetes patient database and analysing it using multiple diabetes characteristics to forecast diabetes illness.

Tiwari et al. [9] used six different machine learning algorithms indicating which algorithm is best algorithm for diabetes prediction. Researchers are interested in diabetes prediction in order to train a software to determine if a patient is diabetic or not by using a suitable classifier on a data set. The categorization procedure, according to prior studies, has not much improved. As diabetes prediction is an important topic in computers, a system is necessary to tackle the difficulties highlighted based on past research.

We examined actual diagnostic medical data based on numerous risk variables in order to classify machine learning techniques and forecast diabetes mellitus in this study.

3 Various Machine Learning Approaches

To predict diabetes, we employ a variety of classification methods. This is a crucial characteristic that plays a big part in prediction. The methods are as follows:

Logistic Regression: In 1958, statistician DR Cox invented logistic regression, which predates the area of machine learning. It is a type of supervised machine learning approach used in classification tasks (for predictions based on training data). Logistic regression employs the same equation as linear regression; however, the result is a categorical variable in logistic regression; whereas, it is a value in other regression models. The independent variables can be used to predict binary outcomes. The logistic regression model is a form of machine learning classification model that has the binary values like 0 or 1, −1 or 1 and true or false as dependent variable and the independent variable like interval, ordinal, binominal or ratio-level [10]. The logistic/sigmoid equation function is as follows:

$$y = \frac{1}{1 + e^{-x}} \tag{1}$$

K-nearest neighbour: Because the data supplied to it is labelled, k-nearest neighbour is a supervised machine learning method. It is a nonparametric approach since the categorization of test data points is based on the training data points closest to it rather than the data set's dimensions (parameters).

Algorithm:

(1) Collect the data set and generate the training and testing sample from it.
(2) Choose an instance from the testing sets and determine how far it is from the training set.
(3) Ascend the list of distances.

(4) The instance's class is the most frequent class among the three initial training examples ($k = 3$).

Support Vector Machine (SVM) Classifier: This SVM classifier is a supervised learning process that excels in pattern identification and is used as a training process for deducing classification and regression rules from data. When the number of characteristics and instances is large, SVM is the most exact method [11]. The SVM algorithm creates a binary classifier. In a SVM model, each data item is represented as a point in an n-dimensional space, with n being the number of features, and each feature as the value of a coordinate in the n-dimensional space. The main goal of SVM is to categorize the data point in a multidimensional space using an appropriate hyperplane. A hyperplane is considered as a boundary of classification for data points. The hyperplane classifies the data points with the biggest gap between the classes and the hyperplane. In this technique, each data item in n-dimensional space is represented as a point, with the value of each feature matching to the value of a certain coordinate.

Random Forest (RF): The random forest is made up of several separate decision trees that operate together as an ensemble, as the name indicates. For each tree, the random forest predicts a class, and the class with the majority of votes becomes our model's prediction report. Random forest is another frequently used supervised machine learning technique [12]. This method works well for both regression and classification issues, although it excels at the latter. As the name indicates, the random forest technique analyses many decision trees before providing an output. As a consequence, it is a decision tree collection. This technique is based on the assumption that if additional trees were present, they would all reach the same conclusion. For classification, it utilizes a voting technique and then selects the class, but for regression, it takes the mean of all decision tree outputs. It works well with large data sets with several dimensions.

4 Methodology

In this section, we will look at the various machines learning classification techniques used for prediction of diabetes mellitus. We will also go through our recommended technique for increasing accuracy. In this article, five alternative techniques were employed. The many methods utilized are listed below. The accuracy measurements of the machine learning models are the output. The model may then be utilized to make predictions.

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome	Id
1991	1	109	56	21	135	25.2	0.833	23	0	1978
1992	2	88	74	19	53	29.0	0.229	22	0	1979
1993	17	163	72	41	114	40.9	0.817	47	1	1980
1994	0	114	80	34	285	44.2	0.167	27	0	1983
1995	4	120	68	0	0	29.6	0.709	34	0	1988
1996	4	110	66	0	0	31.9	0.471	29	0	1989
1997	1	79	60	42	48	43.5	0.678	23	0	1994
1998	8	179	72	42	130	32.7	0.719	36	1	1996
1999	0	129	110	46	130	67.1	0.319	26	1	1998

Fig. 1 Screenshot of few rows of diabetes data set

4.1 Data set Description

The diabetes data set is downloaded from Kaggle repository [13]. A diabetes data set of 2000 cases was used, out of which 684 are affected by diabetes and rest 1316 are non-diabetic. The diabetes data set was used to test the techniques. The goal is to determine whether or not the patient is diabetic based on the measurements. Parameters used in data sets are: (1) age (2) glucose (3) blood pressure (4) BMI (5) insulin (6) skin thickness (7) DiabetesPedigreeFunction (8) pregnancies (9) outcome and (10) ID. Figure 1 depicts few rows of Kaggle diabetes data set.

4.2 Implementation and Design

The study's implementation was done with Google Colab, and the coding was done with the Python programming language. The diabetes data set was used to forecast availability of diabetes using various machine learning approaches such as SVM, k-nearest neighbour and RF and LR classifications. After then, each classifier's predictions are compared to one another. The procedures to implement the machine learning algorithm are illustrated in Fig. 2.

The diabetes parameters serve as the variable which is dependent; whereas, the other factors served as independent ones. For the dependent diabetes features only two values are accepted, with a "zero" indicating no diabetes and a "one" signifying availability of diabetes. The whole sample is divided into two groups, with a ratio of 70:30 for the training and testing data set. All four methods of classification, i.e. were used for prediction. The training data was then used to predict the test set outcomes using SVM, k-nearest neighbour and RF and LR classifications, resulting in the confusion matrix given in Table 1.

The measure provided in Eq. 2–8 may be computed using the obtained confusion matrices. True negative (TN), false negative (FN), true positive (TP) and false positive

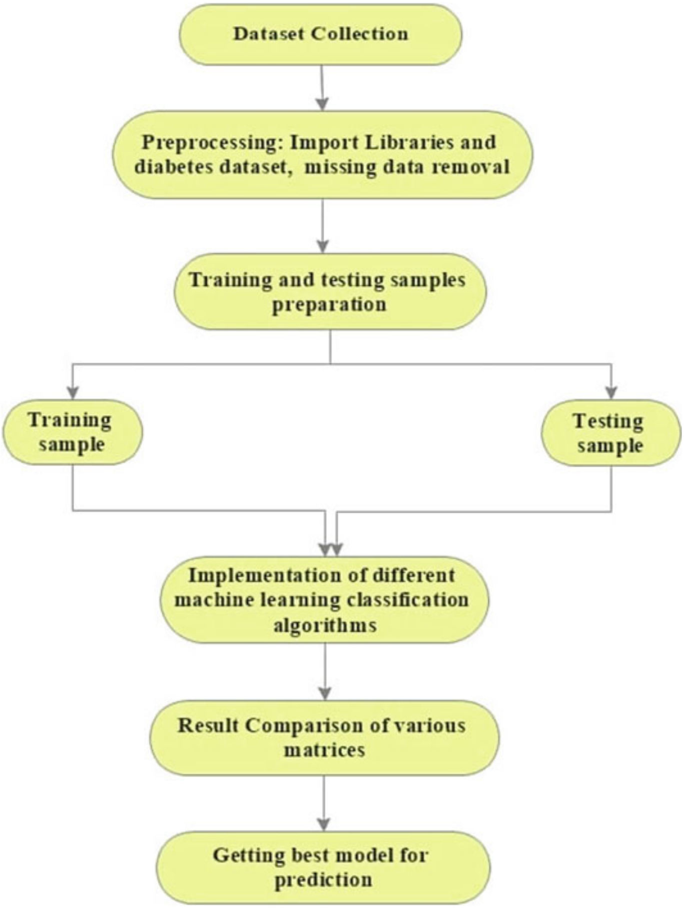


Fig. 2 Proposed flow graph

Table 1 Matrix of confusion for different classification methods

	Logistic regression	K-nearest neighbour	Support vector machine	Random forest
Diabetes data set	[[387 29] [80 104]]	[[336 49] [73 112]]	[[369 21] [119 91]]	[[346 4] [5 245]]

(FP) were the results of these matrices (TP). Because there are more non-diabetic cases than diabetic cases in both data sets, the TN is greater than the TP. As a consequence, all of the techniques provide positive results.

The following measurements have been calculated using the following formulae in order to determine the precise accuracy of each method:

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

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

$$\text{Specificity} = \frac{\text{TN}}{\text{TP} + \text{FP}} \tag{4}$$

$$\text{MCC} = \frac{(\text{TP} * \text{TN}) - (\text{FP} + \text{FN})}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}} \tag{5}$$

$$\text{ErrorRate} = \frac{\text{FN} + \text{FP}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \tag{6}$$

$$F\text{-Measure} = \frac{2 * (\text{Precision} * \text{Sensitivity})}{\text{Precision} + \text{Sensitivity}} \tag{7}$$

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \tag{8}$$

Another finding as per Table 2 is that the accuracy level of all the techniques is higher on our collected data set than on the used diabetes data set, owing to the former's greater number of variables relevant to assessing diabetes risk. The random forest classifier outperforms all others in terms of accuracy (98.4%), sensitivity, specificity, precision and F-measure, proving that it is the best technique for our data set. Furthermore, in the case of random forest, the AUC value is 1, indicating that this model performs exceptionally well in classification. Figure 3a depicts the clear graph for the ROC curve and AUC for the diabetes data set Fig. 3b depicts the comparison of accuracy of all the four machine learning algorithms. Here it indicates

Table 2 Comparison of statistical measurement for various classification techniques

	Logistic regression	K-nearest neighbour	Support vector machine	Random forest
Accuracy	0.818	0.796	0.744	0.984
Error	0.156	0.204	0.255	0.016
Sensitivity	0.775	0.748	0.775	0.987
Specificity	0.666	0.603	0.666	0.916
Precision	0.856	0.832	0.856	0.991
F-measure	0.813	0.787	0.813	0.989
MCC	0.416	0.331	0.416	0.963
Kappa	0.470	0.419	0.466	0.922
AUC	0.865	0.815	0.771	1.000

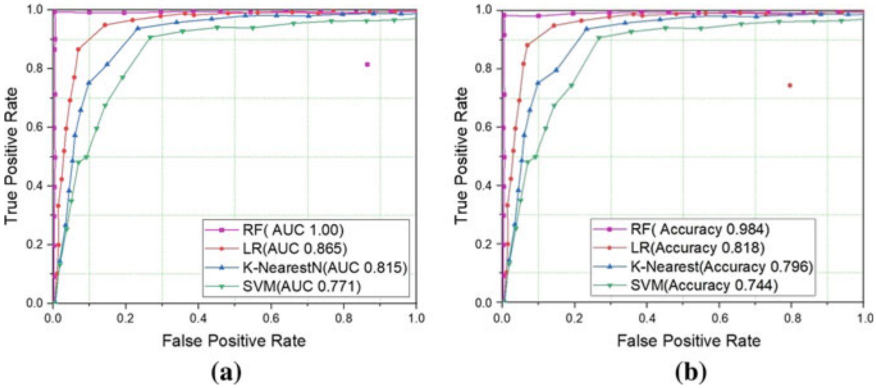


Fig. 3 **a** ROC Curve with AUC for PIMA data set **b** accuracy curve

that in both the cases the RF classifier gives the highest result with AUC value 1 and accuracy of 98.4.

The cross-validation process was also used to assess the efficacy of various models. A subset of the data is set aside for cross-validation, and the remaining data is used to train it. And the procedure is repeated for each segment of data. The size of the pieces is determined by the value of k. Here for validation point of view tenfold cross-validation was used, which means the data was split into ten parts. Cross-validation has the greatest accuracy for random forest. Random forest has a Kappa statistic of better than 0.9, indicating that it is outstanding.

5 Conclusion

One of the most pressing worldwide health concerns is detecting diabetes risk at an early stage. Here in our research that aims to build up a system for predicting the risk of diabetes mellitus. Four machine learning techniques for classification algorithms were used in this work, and the results were compared to several statistical metrics. The above said four algorithms were used on the PIMA database. The accuracy level of RF classification in our data set is 98.4%, which is the greatest among the others, according to the testing results. All of the models generated good results for various parameters like as accuracy, recall sensitivity and so on, using four different machine learning methods. This result can be used to forecast any other illness in future. This study is currently researching and improving on various machine learning approaches to forecast diabetes or any other condition.

References

1. Dash S, Gantayat PK, Das RK (2021) Blockchain technology in healthcare: opportunities and challenges. In: Panda SK, Jena AK, Swain SK, Satapathy SC (eds) Blockchain technology: applications and challenges. Intelligent systems reference library, vol 203. Springer, Cham. https://doi.org/10.1007/978-3-030-69395-4_6
2. Md. Faisal Faruque A, Sarker IH (2019) Performance analysis of machine learning techniques to predict diabetes mellitus. In: 2019 International conference on electrical, computer and communication engineering (ECCE), pp 1–4. https://doi.org/10.1109/ECACE.2019.8679365
3. Chaki J, Thillai Ganesh S, Cidham SK, Ananda Theertan S (2020) Machine learning and artificial intelligence based diabetes mellitus detection and self-management: a systematic review. J King Saud Univ- Comput Inf Sci ISSN 1319-1578. https://doi.org/10.1016/j.jksuci.2020.06.013
4. Alam TM, Iqbal MA, Ali Y, Wahab A, Ijaz S, Baig TI, Hussain A, Malik MA, Raza MM, Ibrar S, Abbas Z (2019) A model for early prediction of diabetes. Inf Med Unlocked 16:100204. ISSN 2352-9148.https://doi.org/10.1016/j.imu.2019.100204
5. VijiyaKumar K, Lavanya B, Nirmala I, Caroline SS (2019) Random forest algorithm for the prediction of diabetes. In: 2019 IEEE international conference on system, computation, automation and networking (ICSCAN), pp 1–5. https://doi.org/10.1109/ICSCAN.2019.8878802
6. Nnamoko N, Hussain A, England D (2018) Predicting diabetes onset: an ensemble supervised learning approach. In: 2018 IEEE Congress on evolutionary computation (CEC), pp 1–7.https://doi.org/10.1109/CEC.2018.8477663
7. Joshi TN, Chawan PM (2018) Diabetes prediction using machine learning techniques. Int J Eng Res Appl 8(1)(Part-II): 09–13. https://doi.org/10.9790/9622-0801020913
8. Rashid Abdulqadir H, Mohsin Abdulazeez A, Assad Zebari D (2021) Data mining classification techniques for diabetes prediction. Qubahan Acad J 1(2):125–133. https://doi.org/10.48161/qaj.v1n2a55
9. Tiwari P, Singh V (2021) Diabetes disease prediction using significant attribute selection and classification approach. J Phys: Conf Ser 1714. (2nd International conference on smart and intelligent learning for information optimization (CONSILIO) 2020 pp 24–25) https://doi.org/10.1088/1742-6596/1714/1/012013
10. Zhu C, Idemudia CU, Feng W (2019) Improved logistic regression model for diabetes prediction by integrating PCA and K-means techniques. Inf Med Unlocked 17:100179. ISSN 2352-9148.https://doi.org/10.1016/j.imu.2019.100179
11. Dash S, Gantayat PK (2020) Liver disease prediction using machine learning algorithm. In: Data engineering and intelligent computing, Proceedings of ICICC 2020. https://doi.org/10.1007/978-981-16-0171-2
12. Yang H, Luo Y, Ren X, Wu M, He X, Peng B, Deng K, Yan D, Tang H, Lin H (2021) Risk prediction of diabetes: big data mining with fusion of multifarious physical examination indicators. Inf Fus 75:140–149, ISSN 1566-2535. https://doi.org/10.1016/j.inffus.2021.02.015
13. https://www.kaggle.com/c/diabetes/data. Data set of diabetes, taken from the hospital Frankfurt, Germany

IoT and RFID: Make Life Easier and Shake up E-commerce Processes with Smart Objects

Shili Mohamed, Kaouthar Sethom, Ahmed J. Obaid, Salwa Mohammed Nejrs, and Saif Al-din M. Najim

Abstract In the field of technological development, Internet of Things (IoT) presents diverse mechanisms to companies to expand their business transactions with customers anywhere and anytime. Processes of selling and buying are getting to be much easier using IoT; therefore, IoT will change the customer's ideas and help them choose the appropriate products. This research article presents the advantage of the Internet of Things and RFID technology in the E-commerce system. From the last year, RFID technology becomes one of the few methods for managing products to guide and help the customer in buying the foremost suitable product which answers his need. This paper explains the use of IoT-supported architecture for E-Commerce with the goal to automate the sale transaction process with physical interaction. It solves problems that companies and customers face in shopping, as knowing our customer desire or knowing which the successful products within the supermarket are.

Keywords RFID technology · E-commerce · IoT · Buyer/seller · Applications

S. Mohamed (✉) · K. Sethom
Innov'COM at Sup'Com Laboratory, University of Carthage, Carthage, Tunisia
e-mail: shilimohamed3@gmail.com

A. J. Obaid
Faculty of Computer Science and Mathematics, University of Kufa, Kufa, Iraq
e-mail: ahmedj.aljanaby@uokufa.edu.iq

S. M. Nejrs
Directorate of Private University Education, Ministry of Higher Education, Baghdad, Iraq
e-mail: salwamn@uomisan.edu.iq

S. A. M. Najim
College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq
e-mail: saifaddin.r@uoanbar.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_13

1 Introduction

E-commerce is a very strong growth field. It encompasses all commercial transactions carried out remotely through electronic and digital interfaces [1]. E-commerce has become the main channel for distance selling; building an online sale is much easier than opening a traditional business. With the Internet there is no limit on location, the marketing of products in physical places is quite complicated and restrictive, but having your E-commerce site is a game-changer, you can sell wherever you want, use the various online marketing tools. Be open 24 h a day and make personalized promotions, the commercialization of smart objects can make life easier [2]. The involvement of connected objects in E-commerce is the most moving novelty since the appearance of the Internet. Indeed, the Internet of Things, first, is a concept and not a specific technology or device. It is the desire to extend the Internet network, and therefore, data exchanges to objects in the physical world, and these objects connected to the Internet can take the shape of an everyday object (e.g., watch, car, TV, glasses, etc.), [3] as the IoT which is growing rapidly. The list of connected objects is very extensive and continues to grow. A connected object is capable of communicating and transmitting information with other connected objects. Hence, we are talking about connected objects that are as intelligent as machine-to-machine (M2M) communication is without the need for humans as an intermediate [4]. Today the Internet of Things constitutes a bridge between the physical world, the world, and the virtual world. In 1990, "300,000" computers were connected to the Internet, in 2000 more than 300,000,000 [5].

Then in 2016, more than 2,000,000,000 smartphones were connected to the Internet. Today, objects are also connected to the Internet and in 2020, we can say that 1.097. 700,000,000 objects are connected. Among these objects, we identify 3.5 billion objects connected to the Internet for navigation systems and other automotive accessories, or 411 billion portable accessories such as connected watches, shoes, glasses, etc. [6]. There are also 646 billion 'objects in hospitals, 9.7 billion are connected to smart city, and 9.5 billion connected to domestic devices. Thanks to a more connected economy, many processes are facilitated. The turnover of the Internet of Things has quadrupled between 2016 and 2020 as shown in Fig. 1. Many sectors have taken advantage of providing the Internet of Things for businesses, which can improve customer satisfaction [7]. Connected objects have several advantages in E-commerce such as the constant personalization of customer recommendations, by collecting data sent via connected objects to your E-commerce where the customer has an account it has also the delivery level orders through cars without drivers and robots [8, 9]. Yet this does not prevent the existence of the inconveniences of the Internet of Things, mainly we have the security problems of personal data. Indeed connected objects produce large quantities of data and the processing of this mass of data involves new concerns, particularly around data confidentiality and security. Connected objects can be hacked like all computer systems [10]. Nowadays, the populations are more and more connected. It demands more and more daily interconnectivity of objects.

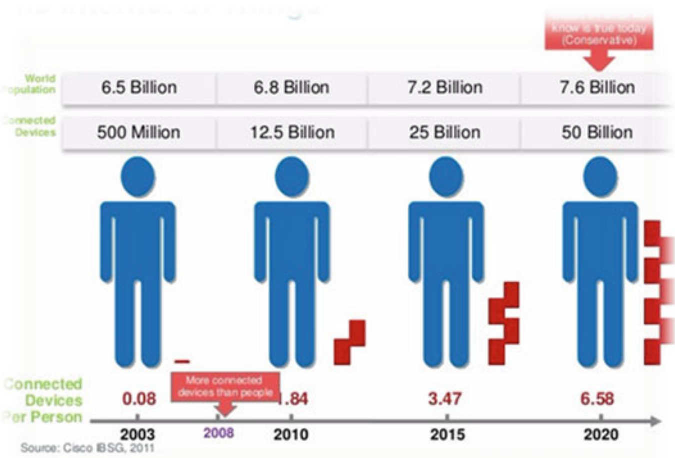


Fig. 1 Internet of Things (IoT) associated devices from 2015 to 2020

The contribution of this paper is to present a sketch of a system to improve sales, by using RFID tags in products and give feedback to the consumers on the point-of-sale. Processes of selling and buying are getting to be much easier using IoT; therefore, IoT is really changing. As well the paper aims at showing new shows a way to influence customers' shopping behavior, through to guide and help the customer in buying the foremost suitable product which answers his need.

The customer became confused when purchasing products; therefore, it is necessary to take into account the requirements of the customer, in response to customer requirements to avoid troubles. We merged IoT and RFID technology are becoming in order to direct and assist customers in purchasing the most appropriate product for their needs.

This paper is organized as follows. Section 1 describes introductory concepts that include E-commerce and the Internet of Things (IoT). In Sect. 2, we introduce the IoT applications and RFID technology. In Sect. 3, we move present the articles dealing with IoT and RFID in E-commerce. Then we expose the problematic in Sect. 4. In next section, we present the case study and thus the proposed system. In Sect. 6, we give a detailed results. Section 7 will be devoted to the discussion. The penultimate section is a presentation of the acronyms and the last will be the conclusion.

2 Background

This section introduces RFID and IoT technologies, as well as, the uses and challenges of RFID technology in IoT.

Fig. 2 IoT (Internet of Things) fields

2.1 IoT Applications

The IoT has the capability to connect every single new object. It has introduced several applications and smart services, which affected users in their daily lives. This section presents a brief discussion of some of IoT applications, which was demonstrated in [11]. It has been introduced to many fields such as social media, health, transport, security, and surveillance as shown in Fig. 2.

2.2 RFID Technology

RFID Tags: They are a bind of microprocessor and they depend on two principle fixings, one of them is a silicon chip, which contains data about things and the other one is a reception apparatus that gets and imparts signs. RFID innovation has the qualities of a long understanding distance, solid entrance capacity, against contamination, high effectiveness, and a lot of data. The RFID framework is principally made out of equipment segments and programming segments. The equipment segments essentially incorporate a peruser and RFID labels it was exhibited in [12] (Fig. 3).

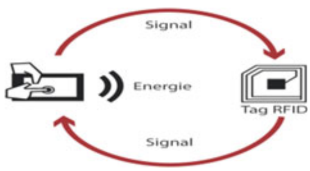


Fig. 3 RFID Tag

3 Related Work

The research methods were suggested by [13]. The study's key aim was to recognize benefits and risks to society as a whole in Industry 4.0, with a focus on customization, which causes many improvements in how markets work (including changes for producers and employees). In the analysis, the three-segment method used in the literature section was limited to the market segment. The literature synthesis and survey methodology "Industry 4.0—perception and aspirations" were used to accomplish this aim. The authors presented a selection of their 2019 research findings. Where the CAWI system was used to perform the research (standardized computer-based Internet interview). A questionnaire with 25 questions was used as the testing instrument (closed, complex, filtering, conditional, and tabular). There were three sections and details of the sample questionnaire. The first section dealt with customization issues, the second with reservations regarding the application of the Industry 4.0 concept in Poland, and the third with the advantages of Industry 4.0.

The authors [14] suggested a new method for recommending products in E-commerce systems that incorporate ontology-based models with the ease of identifying users that have common interests as active users. This method yields more information about the involved user, their respective neighbors, and the products that have been suggested to them (the user), as well as the relationship between them. Our hybrid ontology-based framework can recommend items that fit user expectations more effectively and belong to new categories for the active user compared to the collective filtering version. The suggested method also selects the successful user's k-nearest neighbors, which translates to a limited number of items that can use K-NNAs opposed to the shared filtering variant, which contains a vast range of items to which K-NN will be added, but most of them belong to categories that the active user already recognizes, it meets the active user's expectations. In contrast to the collaborative filtering variant, the experimental evaluation also reveals that the hybrid ontology-based recommender will recommend items to the active consumer that have a higher average overall ranking. Despite the positive effects of the hybrid ontology-based recommendation scheme, the time required to use the K-NN to locate the k-nearest products is excessive. To minimize the time it takes to apply the K-NN algorithm to find the k-nearest products, we intend to find the five most common terms used in all of the product feedback and store this attribute in the database, which will drastically reduce the time it takes to apply the K-NN algorithm to the products. While the system will usually determine the gender of active users based on their names, it would be useful to find a way to determine their gender when the reviewer's name is a pseudonym or is based on numbers. This technique has the potential to improve the consistency of the materials recommended; the proposed approach is shown in Fig. 4.

Given China's vast E-commerce market, the authors [15–18] propose that this study make good use of the massive amount of online consumer comment data to explore and evaluate the dimensions of customer-perceived value and the relevance of each dimension. To gather online feedback data from online reviewers, we created

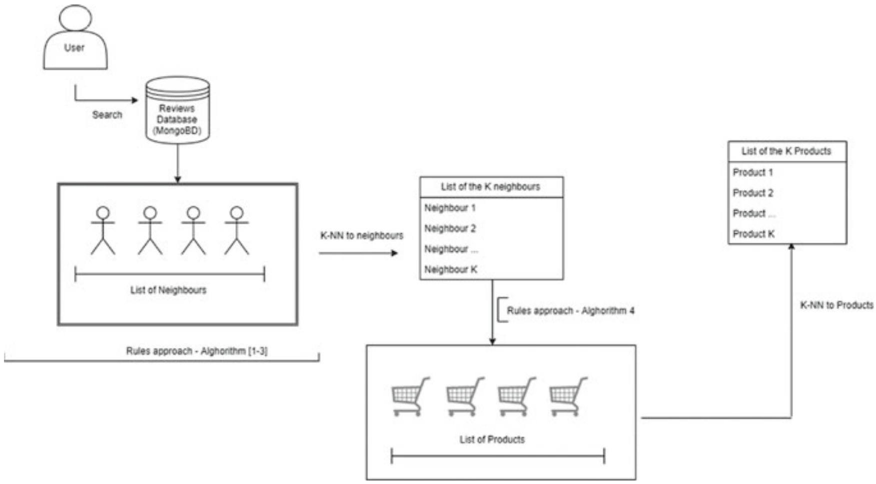


Fig. 4 Proposed approach K-NN algorithm [14]

a Web crawler app. Content processing was conducted on the crawled data after it was preprocessed. The extraction of common words, literature analysis, and expert views are used to create a customer-perceived value dictionary. The measurements of customer-perceived worth were re-identified to incorporate four main dimensions and subdivisions. The dimension model of customer-perceived meaning was tested and implemented for its rationality and operability. The dictionary of customer-perceived meaning is focused on the extraction of often-used words, literature analysis, and expert opinions. For us, the scope of the first-level measurements was defined through literature study. High-frequency terms were automatically extracted, and irrelevant noise was manually removed. The high-frequency terms were then mechanically articulated, generalized, and categorized one-by-one into more standardized and recognizable third-level dimensions. Similarly, the third-level dimensions were more simplified and classified into the more generalized dimensions of the second level to get closer to the first-level dimensions.

The authors [16–19] propose an IoT-based smart e-marketing scheme. Marketers can fulfill consumers' needs for goods and services by using this system and then reach a high level of satisfaction. It also helps marketers to enter new data sources and learn about customer tastes and behaviors. As a result, marketers are better able to provide consumers with what they want. Customers and businesses can quickly fix all of their issues with online shopping by using the proposed method. It also aids buyers, vendors, affiliates, businesses, and advertisers in achieving their objectives with high precision. We used food items as an example of goods in this scheme, but it has the ability to be extended to a wide range of other products in our everyday lives. While the proposed system should deter various attacks and hacking processes, further government involvement in creating a legal foundation is important and would strengthen the proposed system. Marketers can collect a large amount of data about consumers by using the proposed system. This information is stored in our cloud database of customers and includes customer position, time of purchase, a list of sales, and customer demographics.

4 Problematic

IoT takes this universe of commerce business to a totally new level. Shopping in the smartest store can be more complex than it seems, imagine yourself in a smart store doing your shopping, in the range of products diversity, how to choose your products? What are your references? And how can the Internet of Things facilitate this act?

5 Case Study and Proposed Solutions

To answer our problem, we must first define the concepts and issues related to it, and then specify the Spatio-temporal framework of observation. Finally, the presentation includes the resulting of the work of reflection. A final part will allow, through prospective scenarios, to try to imagine the future of connected objects. As these IoT-empowered gadgets trade information over the Web and help computerize our lives, the comfort and productivity of the buying and selling business will undoubtedly improve. The commercial function is mostly affected by technological and IT innovations, being situated in this context, we want to focus on the evolution of the development of electronic commerce, in this context, faced with this problem we offer a solution to facilitate the choice of the brand of the product purchased as follows: First, connect all the smart store products with a unique identifier according to the brand. Secondly, connect the carriages using a specific sensor. When purchasing, the customer passes the product in front of the sensor located at the level of the cart, the latter will display a color either: red, yellow, or green Red: If the purchase of this brand does not exceed 20% of the total purchase of different brands for the same range of products. Yellow: If the purchase of this brand is between 20% and 60% of the total purchase of different brands for the same range of products. Green: If the purchase of this brand exceeds 60% of the total purchase of different brands for the same product line. This can constitute a form on which the customer bases himself to make his choice are presented in Fig. 5.

Saying the existent quantity of an article «*m*» and «*n*» is the number of the bought product of this article, as characterized in Equations.

$$\text{if} \quad n \in \frac{20 * \mathrm{m}}{100} \tag{1}$$

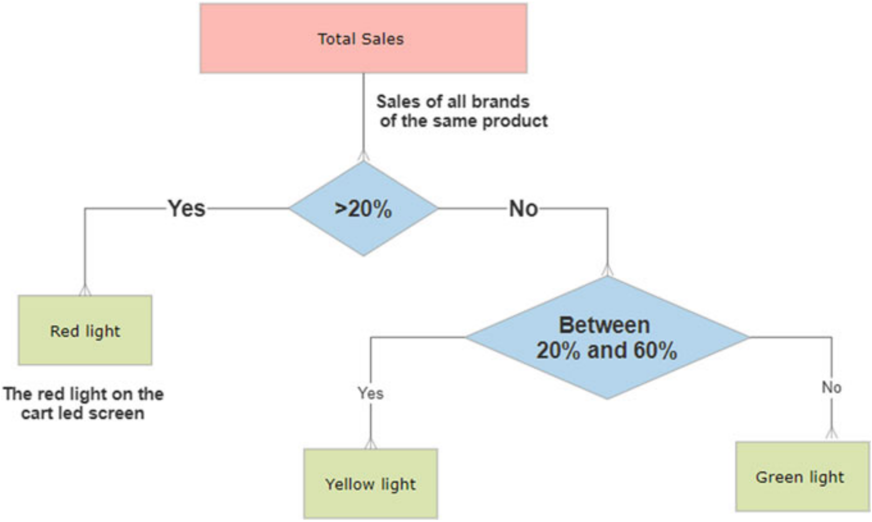


Fig. 5 Product selection diagram

$$\text{if} \quad n \in \frac{20 * m}{100}; \frac{60 * \mathrm{m}}{100} \tag{2}$$

$$\text{if} \quad n > \frac{60 * \mathrm{m}}{100} \tag{3}$$

(1) Red lamp, (2) Yellow lamp, and (3) Green lamp.

P20: Is the probability of a customer who bought 20% of quantity, is as given in Eq. (4) beneath.

P60: Is the probability of a customer who bought 60% of quantity, is as given in Eq. (4) beneath.

$$P_{20} = \frac{1}{0.2 * \mathrm{m}^2} P_{60} = \frac{1}{0.6 * \mathrm{m}^2} \tag{4}$$

After having the case study, we present the proposed system using IOT, as shown in Fig. 6.

Figure 7 depicts the proposed architecture for integrating IoT sensor devices (RFID and Camera), cloud computing (gateway, data center, and Amazon), and the Web application.

Internet of Things has three layers, namely the perception, network, and application layers.

The perception layer is the physical layer, which contains sensors for sensing and collecting environmental data. It detects the customer's profile and the last product he purchased.

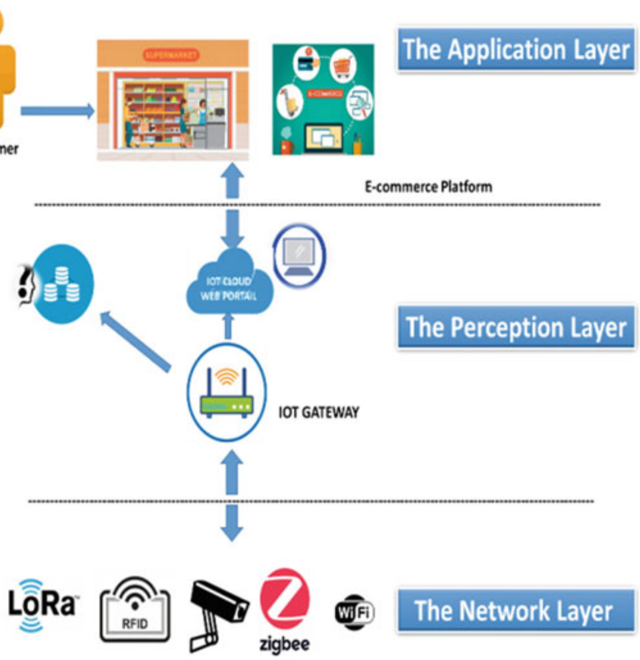


Fig. 6 IoT E-commerce system architecture

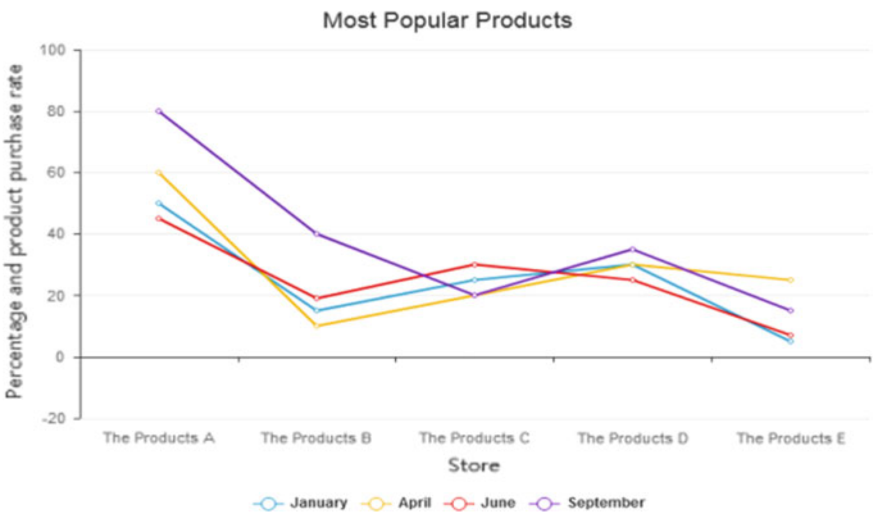


Fig. 7 Most popular products

The network layer is in charge of establishing connections with other smart devices, network devices, and servers. Its capabilities are also used in the transmission and processing of sensor data (Every product is connected tags RFID).

The application layer is in charge of providing the customer with application-specific services. It specifies the number of applications in which the Internet of Things can be used, such as smart supermarkets.

6 Results

This diagram illustrates (Fig. 7) the statistics of the products sold in the store for a period of time, in order that the vendor uses it to understand the requirements of the purchasers and most popular products (These data are according Jumia).

The information was gathered from an Internet business general store that uses the customized item suggestions motor. Furthermore, the "guests who saw this item likewise saw" item proposal creates the most income. To succeed in E-commerce, you would like three things: wanted items to supply, the skills to advertise them, and therefore the drive to succeed. You as of now have the victor's mentality—you are ceaselessly learning, investigating, and preparing for progress. In any case, finding documented items to sell are often a test. That is the rationale we update this rundown consistently to help you with revealing moving item thoughts. In this rundown, you will find moving items in 2021 which will enact new business thoughts—or, you will locate an unprecedented item to feature to your current store. Since each moving item is meant for a specific crowd, we are additionally sharing specialty explicit tips for arriving at new clients. The things on this rundown are the highest moving item classifications from 01 January 2020–31 December 2020 on the Jumia stage, the example as shown in Fig. 8. Jumia is phenomenally contrasted with other

Fig. 8 Customer satisfaction score

Fig. 9 Assisting customers in selecting the most appropriate product

known online business objections in Tunisia which makes shopping easy to try to at completely moderate expenses.

IoT-empowered gadgets to be able to refine client commitment through personality-based data, and considerably more. Subtleties like this make it a lot simpler to customize promoting to suit a particular client. A particularly customized client assistance experience upgrades client commitment. IoT gadgets assist sellers with making a more refined and personalized E-commerce experience for every customer, which thus incredibly improves changes. The diagram illustrates a prototype a customer in a store, shown in Fig. 9.

7 Discussion and Future Recommendation

By 2030, it is anticipated that IoT will be fully operational, with smart sensors interacting with organization servers to help in device personalization. Sensors will automatically detect which objects are nearing completion, and they will send this information to the enterprise servers. After reviewing the user's previous experiences and preferences, the website will make suggestions to the user that are suitable to their tastes. A client's purchase of a product will be suggested to the client's friends via social media. Present methods will not be able to handle the massive amounts of records, IP addresses, and consumer traffic in the coming era of "IoT" and "Big Data." For science, new methods focused on Big Data and IoT would be needed. The proposed methodology makes use of IoT and can therefore accommodate a large amount of data. New E-commerce architecture and adaptation strategies are needed to align with existing E-commerce application platforms, thus increasing the accessibility of Web and wireless networks and improving the mobile user experience. For E-commerce networks, m-commerce connectivity services pose a new threat.

8 Conclusion

On the real level, many objects are now connected to the Internet; the population is getting surrounded by interactive objects. Thanks to artificial intelligence, everything is smart such as a smart watch, smart city, smartphone, smart agriculture, and we have smart commerce. In the first place, the IoT is used in commerce to reduce costs, mainly logistics, and increase productivity, which influences customer satisfaction. Let us quote the example of the out of stocks, which are inevitable. It is like a disaster mainly when there are promotions advertised on TV and radio and you are going to bring in the customer, during this period, the sales will be accelerated and multiplied and it will be complicated to report constantly and be present then that the IoT offers robots that will scan the shelves and send notifications to workers. This technique will speed up the task and know where to report as a priority. In a second place, the IoT was used from a customer perspective, which must be characterized, by its usefulness and ease of use. These innovations must make the life of the customer more with the least cost, short deadlines, and hypersatisfaction for the consumers. The IoT is will be is no longer a novelty and tends to become the norm, the whole world thus becomes one and only one information system.

Finally, as other perspectives of the IoT in E-Commerce, we can mention the Web Service (REST, SOAP) and ontology as advantages.

References

1. Hitpass B, Astudillo H (2019, January) Editorial: industry 4.0 challenges for business process management and electronic- commerce. J Theor Appl Electron Commerce Res
2. Said HM, Salem A-BM (2019) Smart E-business model based on block chain (BC) and internet of things (IoT) technologies. Int J Internet Things Web Services 4
3. Bena YA, Gulumbe MA, Abd. Rahman A (2017, February) IoT on E-Commerce, present and future: a review of Alibaba case study. J Inf Syst Res Innov 11(1):41–46
4. Atlam HF, Walters RJ, Wills GB (2018, September) Internet of things: state-of-the-art, challenges, applications, and open issues. Int J Intell Comput Res (IJICR) 9(3) (Copyright © 2018, Infonomics Society 928)
5. Jayaram A (2017) Smart retail 4.0 IoT consumer retailer model for retail intelligence and strategic marketing of in-store products. In: Proceedings of the 17th international business Horizon-INBUSH ERA-2017, Noida, India, 9 February 2017
6. Strzelecki A (2019) Key features of e-tailer shops in adaptation to cross-border e-commerce in the EU. Sustainability 11:1589
7. Landaluce H, Arjona L, Perallos A, Falcone F, Angulo I (2020) Florian Muralter a review of IoT sensing applications and challenges using RFID and wireless sensor networks. Sensors 20(9):2495. https://doi.org/10.3390/s20092495

8. Bayer S, Gimpel H, Rau D (2020) IoT-commerce - opportunities for customers through an affordance lens. Electron Mark. https://doi.org/10.1007/s12525-020-00405-8March
9. Hoffman DL, Novak TP (2018) Consumer and object experience in the internet of things: an assemblage theory approach. J Consumer Res 44(6):1178–1204. https://doi.org/10.1093/jcr/ucx105
10. Jing N, Jiang T, Du J, Sugumaran V (2018) Personalized recommendation based on customer preference mining and sentiment assessment from a Chinese e-commerce website. Electron Commerce Res 18(1):159–179
11. Jayaram A (2017, February) Smart retail 4.0 IoT consumer retailer model for retail intelligence and strategic marketing of in-store products. In: Proceedings of the 17th International business Horizon-INBUSH ERA-2017, Noida, India, 9
12. Mekruksavanich S (2020) Supermarket shopping system using RFID as the IoT application. In: 2020 joint international conference on digital arts, media and technology with ECTI Northern section conference on electrical, electronics, computer and telecommunications engineering (ECTI DAMT & NCON). https://doi.org/10.1109/ECTIDAMTNCON48261.2020.9090714
13. Saniuk S, Grabowska S, Gajdzik B (2020) Social expectations and market changes in the context of developing the industry 4.0 concept. Sustainability 12(4):1362. https://doi.org/10.3390/su12041362
14. Saniuk S, Grabowska S, Gajdzik B (2019) A hybrid ontology-based recommendation system in e-commerce. Algorithms 12(11):239. https://doi.org/10.3390/a12110239
15. Liu C, Wang S, Jia G (2020) Exploring e-commerce big data and customer-perceived value: an empirical study on Chinese online customers. Sustainability 12:8649. https://doi.org/10.3390/su12208649
16. Abdel-Basset M, Mohamed M, Chang V, Smarandache F (2019) IoT and its impact on the electronics market: a powerful decision support system for helping customers in choosing the best product. Symmetry 11(5):611. https://doi.org/10.3390/sym11050611
17. Pandey D, Wairya S, Al Mahdawi R, Najim S, Khalaf H, Al Barzinji S, Obaid A (2021) Secret data transmission using advanced steganography and image compression. Int J Nonlinear Anal Appl 12(Special Issue):1243–1257. https://doi.org/10.22075/ijnaa.2021.5635
18. Ibrahim K, Obaid A (2021) Fraud usage detection in internet users based on log data. Int J Nonlinear Anal Appl 12(2):2179–2188. https://doi.org/10.22075/ijnaa.2021.5367
19. Manaa ME, Obaid AJ, Dosh MH (2021) Unsupervised approach for email spam filtering using data mining. EAI Endorsed Trans Energy Web.https://doi.org/10.4108/eai.9-3-2021.168962

Adopting a Blockchain-Based Algorithmic Model for Electronic Healthcare Records (EHR) in Nigeria

Gabriel Terna Ayem, Salu George Thandekkattu, and Narasimha Rao Vajjhala

Abstract This work seeks to explore solutions to the challenges posed in sharing/integrating/transferring EHR data across heterogeneous healthcare institutions in developing nations, particularly in Nigeria. This blockchain-based algorithm model for EHR interoperability seeks to address EHR interoperability challenges such as semantics across heterogenous healthcare institutions, the proper infrastructure and consensus structure for the sharing of EHR across healthcare institutions, privacy and security of patients' records, etc. Hence, this algorithm model, when adopted, is adjusted and contextualized to fit developing nations by addressing the underlining fundamental challenges. EHR interoperability would become a reality across the heterogeneous healthcare institutions in these nations. Future works on EHR will focus on the aggregate blockchain model. Also, the Artificial Intelligence (AI) model inculcating blockchain technology would be an attractive option to dive into future works on EHR.

Keywords Electronic Healthcare Records (EHR) · Electronic Medical Records (EMR) · Healthcare Information Systems (HIS) · Interoperability · EHR challenges · Blockchain · EHR semantics · Security · Privacy

1 Introduction

Information Technology (IT) or Information and Communication Technology (ICT) is the driving force behind innovation and processes in twenty-first-century society. Computer algorithms that transform manual processes and procedures into automated

G. T. Ayem · S. G. Thandekkattu
American University of Nigeria, Yola, Nigeria
e-mail: gabriel.ayem@aun.edu.ng

S. G. Thandekkattu
e-mail: george.thandekkattu@aun.edu.ng

N. R. Vajjhala (✉)
University of New York Tirana, Tirana, Albania
e-mail: narasimharao@unyt.edu.al

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_14

processes in organizations are constantly on the increase. Daily, engines, robots, etc., in industries and organizations are designed, developed, and deployed using the computer algorithm model as their driving force. The term information systems refer to all the processes and procedures involving computing components like hardware, software, telecommunication, people, and procedures.

In the healthcare sector, computers are used intensely to drive care in the industry. Terminologies, such as Electronic Health Records (EHR) system, Healthcare Information Systems (HIS), Electronic Medical Records (EMR), among others, have gradually become the computer terminologies used in the implementation of computerization in the healthcare sector [1, 2]. Developed economies are at the forefront of this computerization of the healthcare sector, albeit the developing economies such as Nigeria are not precluded in this phenomenon, even though not to a greater extent than their developed counterpart. The implementation of EHR is no longer an issue, as it has become a common phenomenon even among developing economies, as the impacts and gains seem plausible [3]. However, when it comes to information or data exchange among two or more healthcare institutions, exchanging or sharing their EHR among themselves to facilitate or enhance individual care becomes a challenge. This is as a result of some impeding factors such as the right model or algorithm to use for EHR interoperability that could properly handle delicate, ethical, and legal issues such as security and privacy of patients' records, semantics among others which pose as major challenges to EHR interoperability even the world over. Other issues such as the right kind of network infrastructure, standards, and the right consensus model that would drive this interoperability process are still daunting tasks. A centralized system, a client–server system, cloud-based system, and bulk transmission infrastructure frameworks have all been proposed as solutions to this issue [4]; these frameworks of interoperability infrastructures have their downsides which borders on the issues of trust, operational control over records or information, and the right model to handle diverse semantics that is employed by different healthcare institutions and their practicing personnel [5, 6].

This study seeks to explore extant works in the areas of EHR and interoperability in healthcare institutions in developing economies of the world, focusing on Nigeria, using the framework proposed by Peterson et al. [6] to contextualize and underpin the work. Peterson et al. [6] work on interoperability using the blockchain algorithm model seems to overcome most of the challenges faced in EHR interoperability, such as trust, security, privacy, semantics, and other sundry challenges in network interoperability of EHR.

2 Review of Literature

2.1 Electronic Health Records (EHR)

Electronic Health Record (EHR) or Electronic Medical Records (EMR) are electronic documents that are unique and composed of patient data or information relating to their health history, medical prescriptions, demographics, laboratory results as well as tomographic reports, etc. [7]. EHR and EMR are at the forefront of Information and Communication Technologies (ICT) application solutions that have been advanced and employed primarily to supervise patients' healthcare records or information [8]. However, both EHR and EMR are considered to be the same and are frequently used interchangeably. There is a thin difference between the two. EMR usage is mainly restricted to or in a healthcare sector only during EHR traverse. It transcends healthcare sectors as patient's information or records could come from diverse sources aside from the healthcare organization, e.g., an EHR may receive or inculcate records into its database from sources like a government repository, etc. [9, 10]. Also, the EHR can provide real-time information centered around the patient, such as history diagnosis, test results, and allergies, that are employed in making proper clinical decisions regarding the patient. A prominent feature in EHR is that vendors and staff who are authorized access can modify patient records anytime and place. Both EMR and EHR systems are collectively called Healthcare Information System (HIS), which are very useful. They are employed in healthcare institutions to help improve healthcare professionals' proficiency in the discharge of their duties, thereby reducing time and costs and bringing about the encouragement of medical practices based on evidence [11].

Developing countries like Nigerian healthcare institutions are faced with a plethora of challenges in the effective implementation of EHR without recourse to compromising the security and privacy of patients' records. Nowadays, many secondary and tertiary healthcare institutions in Nigeria have implemented EHR, EMR, or HIS in their respective hospitals to help manage patient information in a digital format. These preclude compatibility integration as they are usually designed and developed by diverse firms or vendors and are implemented in different high-level languages and databases [12]. This process is unique and very effective when implemented in the same hospital or healthcare institution where this EHR is deployed.

2.2 Interoperability

According to the Electrical Electronics Engineers standard dictionary of computing, interoperability is defined as the ability of two or more components to exchange information and use the information exchanged. A more concise definition of interoperability by Peterson et al. [6] states "the extent to which the clinical intents could be conveyed across institutional healthcare boundaries." There are generally two main acceptable approaches to effective interoperability:

- Syntax: Syntax outlines the structural format on how the information should be traded. The syntax is oblivious of the contents that are being exchanged.
- Semantics: Semantics ensure that both networks engage in the data exchange of EHR and comprehend the information traded across both networks. This happens by incorporating the appropriate metadata; for example, in the taxonomy of medical science, classification means groups or lists of similar objects such as infections and laboratory tests. In contrast, ontology infers hierarchical relationships among concepts such as bacterial culture used to diagnose infections.

2.3 Challenges of EHR Interoperability

1. Privacy and Security of Records: In many nations, the privacy and security of patient records is a legal and ethical matter [6]. Therefore, any healthcare institution that fails to protect and secure its patient's healthcare record, especially when exchanging them across other institutions' networks, may likely face litigations, a ban or withdrawal of services which may lead to a loss in revenue or shortage in finances as a result of low patronage by patients. Hence, this issue can be a significant challenge and a big setback in the wheels of progress in EHR interoperability. The phobia of compromising on it could sometimes outweigh the potential benefits that are concurrent with EHR exchange.
2. Lack of uniform and agreeable architectural framework across Networks: A lack of consensus architectural infrastructure among healthcare institutions and relevant stakeholders is another con that is bedeviling EHR interoperability across institutions' networks. Researchers and stakeholders in the healthcare sector have made several attempts to address this issue. Many have suggested a centralized data source, sending a large volume of data across institutions nodes, etc. These options often pose a unique problem. For example, a centralized approach is likely to increase security breaches, and control is usually shifted to a single centralized authority. The issue of complete trust by all parties' concerns may not be guaranteed. On the contrary, the humongous exchange of data would compel individual healthcare originations to lose control of their operational [6].
3. Heterogeneous and Complex Nature of Healthcare Data: Healthcare data is very complex and highly heterogeneous, thus making EHR interoperability a challenging task to surmount, as different healthcare organizations have diverse semantics and meanings for the same kind of data item. This work seems to proffer suggestions on the best approach of achieving EHR interoperability across different healthcare institutions by adopting the blockchain algorithm combed with FHIR, a standard exchange protocol set by level 7 health. And

we propose to adopt and contextualize it in the Nigerian Healthcare institutions system. This proposed approach seems to tackle the significant issues of structure and semantics bedeviling EHR interoperability [6].

3 Nigerian Healthcare Structure

The Nigeria healthcare sector is a collaboration of the government (both state and federal), the private sectors, and the religious organizations. The policies, technical operations, and running procedures are the responsibility of the federal government of Nigeria. Similarly, in their capacities and mandate, various state governments make policies and procedures to support the healthcare centers that are established and controlled by them. The local government healthcare centers, being the third tier and last on the rung of the healthcare ecosystem and its value chain in Nigeria, are regulated by the state government and financially supported by state and federal governments.

The issues raised as challenges to effective implementation of EHR and EHR data interoperability in Nigeria are as follows: inadequate ICT infrastructure and facilities, lack or inadequate power supply, lack of political will to change extant health customs and policies to reflect best global practices, incompatibility in many e-health system, breech of patient security and privacy, corruption and above all, lack of consensus ontologies, and standard in the healthcare system to handle semantics in areas that require standardization. Their study recommended measures that would correct or mitigate these challenges they identified as impediments to EHR implementation and data interoperability in Nigeria [13].

A framework that would break the impediments to effective EHR implementation and could lead to EHR data interoperability in Nigeria was proposed by Kruse et al. [14]. *Socialized Medicine* the name given to their framework was composed of a communication infrastructure, with a wired or wireless network, a mobile phone, and Internet connectivity. This Internet connectivity will enable doctors and patients to use their smart phone and other devices to communicate virtually and share information across heterogenous healthcare institutions platforms. Albeit for their framework to really be deployed and implemented in real time, the following components and infrastructure would be indispensable: A Web-based application software implementing a cloud computing infrastructure would be needed, a server-side scripting tool, and an integration of a real-time communication tool will also be required. The work proposed authorization and licensing for all category of users to ensure safety in the system. Three key gaps were identified from the review of the literature:

1. Differences in data semantics across healthcare institutions networks.
2. Guaranteed security and audit of patient records across the networks, and
3. Lack of control by patients over their records across the networks.

Although in the three gaps identified above, Pai et al. [15] and Bahga and Madisetti [16] seem to deal with some extent on semantics and security, leaving the issue of patient control on their data. They employed the cloud-based EHR approach alone to decipher security and semantics issues. This kind of architecture is not tested on the traditional client–server architecture to ascertain its workability regarding EHR interoperability in these nations under review. This study intends to close the gaps mentioned earlier by employing the blockchain-based algorithm approach proposed by Peterson et al. [6] to solve the current gaps of EHR interoperability in developing nations with particular reference to Nigeria's healthcare systems.

4 EHR Interoperability Blockchain Algorithm

A blockchain is a distributed transaction ledger, which is made up of blocks such that each block represents a set of transactions [17–19]. The properties of a blockchain data structure are interesting, and the chains of blocks that make up the blockchain are considered immutable. The immutability of a blockchain is achieved as a result of a verification number called a hash present in each of the blocks, which verifies each transaction's genuineness in the block. Also, a current hash block is dependent on a predecessor hash block. Hence with this scenario, it goes to say any transaction change occurring in a particular block hash will translate into changing the transaction history of the previous hash block as well. Thus, this setup makes the blockchain immutable [6, 18, 20, 21]. Blockchain technology precludes a distributed system, as it does not depend on a centralized trusted system for the verification and validation of information or data. It requires the use of another mechanism to reach a consensus before transmission can occur. The PWF makes sure that any node on the network that wishes to add to the network must carry out a complete computationally verifiable puzzle as a matter of requirement.

Similar to the digital currency of Bitcoin, data blocks are added to the system at regular intervals of time. The proof of work is an interval determinant of time in the Bitcoin system. This work specifies a constant interval of time for creating a unit of block or a block period as recommended by Peterson et al. [6]. Accordingly, during this block period, the nodes go through four activity phases. The first phase is the transaction distribution phase; the time commences at Tα, after which transactions are passed over to the coordinating node or the miner node. This phase may continue until Tσ or until the miner node stops accepting a new transaction for the block. The coordinating node will then collect the entire new block and send it to the nodes for review in the block verification request phase, which is the second phase. This enables all nodes which contributed to the block in at least one transaction entry to show their approval by endorsing the block digitally. The block control is then passed to the miners in phase three's signed block return phase. The miners then add up the block to its local blockchain. Finally, distribution to the new blockchain is made in the fourth and final phase, called the new blockchain distribution phase. These phases are illustrated in Fig. 1.

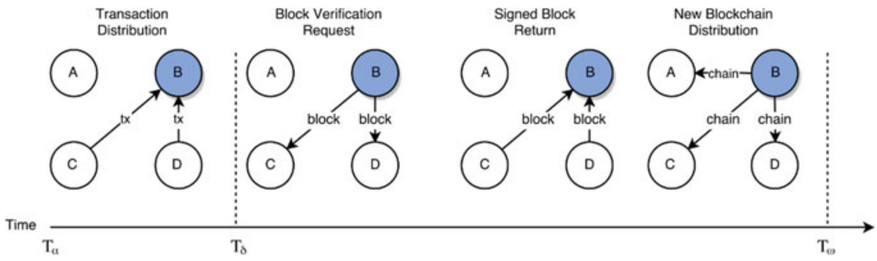


Fig. 1 Shows the four phases of adding a block to the EHR blockchain

- Data Search and Access: As stated earlier, the actual patient's record is not in the block but the FHIR. With this model, searchability is possible and imperative. To perform a search, search privileges to query the block are assigned to appropriate external healthcare institutions, which carry out searches with the help of keywords resident in the secure index file of the transaction.
- Data Security: The blockchain model applied in this EHR work has its strength in the security, privacy, and anonymity of patients' transaction records. The blockchain encryption process described under the data search and access process is employed to secure this model. The network will also encrypt public information and direct all nodes shared key, while the source node should encrypt sensitive information. Finally, a security technique known as the privacy-preserving key (PPK) should be employed to enhance data search and access; thus, if an external network requests a set of record transaction from the blockchain that matches specific criteria, both the query and the transaction will remain encrypted and secure.
- Patient's Identification: The ability of EHR interoperability models to consistently identify patients across different healthcare institutions can be a daunting task. Studies have suggested some variation of the records in the central master plan index (MPI) file. In this model, similar to the blockchain Bitcoin model, data allocation to the model is done using addresses instead of the patient's name. At the same time, the key control to the addresses is assigned to the patient. The pro of this method is manifest in the fact that a single identifier consensus would not be required as a particular patient may have many different healthcare institutions' addresses. The procedure is a great significant departure from extant practices where healthcare institutions implement EHR own and assign patient's identifiers mechanism [6].

5 Conclusion

Interoperability of EHR has been a daunting challenge for developing economies and nations, with Nigeria not precluded in the list. The study revealed that the decentralized nature of the model makes it imperative and germane in handling the plethora of issues of EHR interoperability in the context of the review nations, as it addresses foundational and chronic problems of interoperability of EHR in these nations hitherto were barriers. In this paper, the authors presented topics such as semantics, privacy, search, patient identifier across different networks of healthcare institutions, and network consensus structure. Thus, we consider this model the most suitable and secure EHR interoperability model among the extant models on this subject; as the benefits of this model on both patients and healthcare institution speaks volumes. Future works on EHR will focus on the aggregate blockchain model. Also, the Artificial Intelligence (AI) model inculcating blockchain technology would be an attractive option to dive into EHR.

References

1. Tritilanunt S, Ruaysungnoen S (2016) Security assessment of information system in hospital environment. In: Proceedings of the fifth international conference on network, communication and computing. Association for Computing Machinery: Kyoto, Japan, pp 11–16
2. Wani TA, Mendoza A, Gray K (2019) BYOD in hospitals-security issues and mitigation strategies. In: Proceedings of the Australasian computer science week multiconference. Association for Computing Machinery, Sydney, NSW, Australia, Article 25
3. Braghin S et al (2008) Introducing privacy in a hospital information system. In: Proceedings of the fourth international workshop on software engineering for secure systems. Association for Computing Machinery, Leipzig, Germany, pp 9–16
4. Clim A, Zota RD, Constantinescu R (2019) Data exchanges based on blockchain in m-Health applications. Proc Comput Sci 160:281–288
5. Dauda I et al (2021) Blockchain technology in healthcare systems: applications, methodology, problems, and current trends. ATBU J Sci Technol Educ 9(1):431–443
6. Peterson K et al (2016) A blockchain-based approach to health information exchange networks. In: Proceedings of NIST workshop blockchain healthcare
7. Kontzinos C et al (2020) Combining distributed and centralised approaches to develop a hybrid methodology for data management in healthcare. In: Proceedings of the IADIS international conference on WWW/Internet, pp 35–42
8. Adi S (2021) Psychological effect of brand image and brand reputation on sustainable firm performance in indonesian logistic companies. Ilkogretim Online 20(4):820–831
9. Hydari MZ, Telang R, Marella WM (2015) Electronic health records and patient safety. Commun ACM 58(11):30–32
10. Ebadollahi S et al (2006) Concept-based electronic health records: opportunities and challenges. In: Proceedings of the 14th ACM international conference on multimedia. Association for Computing Machinery, Santa Barbara, CA, USA, pp 997–1006
11. Liao S et al (2017) Towards use of electronic health records: cancer classification. In: Proceedings of the symposium on modeling and simulation in medicine. Society for Computer Simulation International, Virginia Beach, Virginia, Article 4
12. An S et al (2019) Longitudinal adversarial attack on electronic health records data. In: The world wide web conference. Association for Computing Machinery, San Francisco, CA, USA, pp 2558–2564
13. Ijeoma Nkem O-A et al (2021) Identifying advocacy strategies, challenges and opportunities for increasing domestic health policy and health systems research funding in Nigeria: perspectives of researchers and policymakers. Health Res Policy Syst 19(1):1–11
14. Kruse CS et al (2016) Barriers to electronic health record adoption: a systematic literature review. J Med Syst 40(12):252

15. Pai MMM et al (2021) Standard electronic health record (EHR) framework for Indian healthcare system. Health Services and Outcomes Research Methodology
16. Bahga A, Madisetti VK (2013) A cloud-based approach for interoperable electronic health records (EHRs). IEEE J Biomed Health Inform 17(5):894–906
17. Negash S, Thomas D (2019) Teaching blockchain for business. In: 2019 IEEE Canadian conference of electrical and computer engineering (CCECE). IEEE, Edmonton, Canada
18. Rao AR, Dave R (2019) Developing hands-on laboratory exercises for teaching stem students the internet-of-things, cloud computing and blockchain applications. In: 2019 IEEE integrated STEM education conference (ISEC). IEEE, Princeton, NJ, USA
19. Strang KD, Che F, Vajjhala NR (2020) Ideologies and issues for teaching blockchain cybersecurity in management and computer science. Springer
20. Liu X (2018) A small java application for learning blockchain. In: 2018 IEEE 9th annual information technology, electronics and mobile communication conference (IEMCON). IEEE, Vancouver, Canada
21. Kianmehr K, Alhajj R (2008) Effectiveness of support vector machine for crime hotspots prediction. Appl Artif Intell 22(5):433–458

Design and Development of IoT Wearable Device for Early Detection of COVID-19 and Monitoring Through Efficient Data Management Framework in Pre-pandemic Life

M. R. Sundara Kumar, Ahmed J. Obaid, S. Sankar, Digvijay Pandey, and Azmi Shawkat Abdulbaqi

Abstract COVID-19 virus named CORONA is a vigorous disease spread all over the world very quickly and creates a pandemic situation to the human beings normal life. As per the doctors and researchers from the laboratory point of view, it will spread to a huge volume when humans are not followed certain principles. Moreover, this disease is easily transferred to neighbors and others in a short period which leads to death. To rectify the remedy for this virus, various spread countries and research peoples are creating the vaccines and some precautionary methods for living hood situation. Recent techniques are used to detect and monitoring the COVID-19–affected person's lifestyle and insisting they take precaution steps for early pre-pandemic life. IoT is a framework that is used to generate data from the human body from the sensors opted for human conditions. Wearable devices have been created with these sensors and communicated with human bodies directly or indirectly. The generated data will send through the server using any connectivity techniques such as Bluetooth or Wi-Fi. Analytics will be done at the server side for taking actions like the human body is affected by the COVID-19 virus or not. Finally, the generated data from a human can continuously store in real time in a cloud server which will be managed as a framework efficiently. This research work proposes a framework for data management in the early detection and monitoring of COVID-19 persons through IoT wearable devices in a pre-pandemic life. The experiments have been

M. R. Sundara Kumar (✉) · S. Sankar
Department of CSE, Sona College of Technology, Salem, India
e-mail: sundar.infotech@gmaill.com

A. J. Obaid
Faculty of Computer Science and Mathematics, University of Kufa, Kufa, Iraq
e-mail: ahmedj.aljanaby@uokufa.edu.iq

D. Pandey
Department of Technical Education, IET, Dr. A.P.J. Abdul Kalam Technical University, Lucknow, Uttar Pradesh, India

A. S. Abdulbaqi
Department of Computer Science, College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq
e-mail: azmi_msc@uoanbar.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_15

done at different zones, and the results are shown symptoms of COVID-19 disease. Parallel work reveals the data management in a cloud server since data have generated continuously in real time and tracking details also stored genuinely. Data management is the typical process in this research because all the data were generated in real time and analytics will be done whenever required. For that large amount of space and effective retrieval technique is required for data extraction. This research work data set is derived from various Internet sources like government web sites and mobile applications, and then, results have displayed the COVID-19 disease details accurately in real time.

Keywords COVID-19 · Bluetooth · Wi-Fi · IoT · Cloud server · Framework · Wearable device

1 Introduction

COVID-19 virus affected the normal life of a human all over the world. In this digital era, new techniques and innovative ideas are used but early detection or assessment is not possible due to the lack of real-world application problems. In the COVID-19 scenario, it affects the normal person which is having low immunity power and not properly maintained the social distance from the affected persons [1]. COVID-19 created a pandemic situation in all countries of the world and monitoring the individual persons which are not affected by this dangerous virus in the early stage is a challenging one. After pandemic life, it would be slightly easy to monitor the person by keep track of their movements regularly. But in this early stage, it is not possible to track all the human movements and stored them in a commonplace regularly [2]. An IoT-based framework is used to maintain the above-said problem with the help of the sensors which are accessed the human body actions, and the results will incorporate into the framework for the decision-making process. Not all types of sensors are allowed to directly connect with the human body; it may lead to side effects in the body. Limited frequency and functioned sensors are used directly with the human body to generate data from human actions [3]. These will take it as parameters and considered the affected person factors are generated as digital data for storing and analytic purposes. Based on the results given by the analytics from the cloud server, treatments and precautions will be suggested to the non-affected persons [4]. Moreover, the zone-level persona is identified and insisted them to strictly follow the rules for the COVID-19 virus. All over the world, a huge number of peoples is affected COVID-19 virus and it will be spread to the other people elastically within a short-range [5]. All countries are affected by this virus and innovative ideas would be suggested from the scientists along with doctor's society. Figure 1 show the peoples are affected by COVID-19 virus analysis.

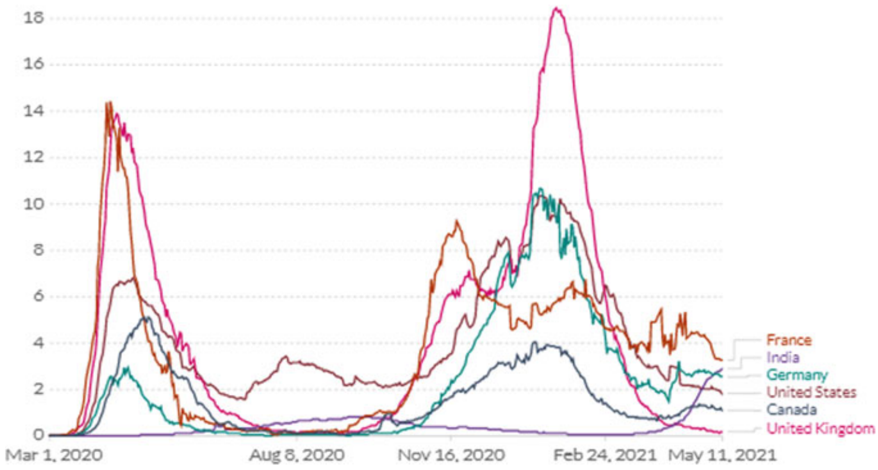


Fig. 1 COVID-19 cases in different countries

2 Related Works

COVID-19 virus research work has been carried out by all the countries in the world since 2020 onwards but the rate of spread and death is not controlled globally. Several innovative ideas revealed from scientists that vaccine is the solution to avoid spreading of the virus but experimentally it would take time to complete the process [6]. Until that, some precautionary methods and actions would have carried out by the peoples in and around the world. Some of the countries found a method to tracking the person's activities in the different sessions and insist they take precautionary actions against corona [7]. Few countries tried to prepare vaccines from natural resources and advised the people to take heavy nutrition foods. A few countries did innovative ideas to track the post-pandemic people's life for avoiding the spread of the virus to their neighbors [8]. Developing countries are worked on a data management framework for storing all kinds of data generated by the peoples in their countries. But the challenges in that work have to manage them with proper techniques and control through a centralized server [9]. The algorithms and protocols used to do that work have been carried out with recent trends like machine learning, deep learning, and IoT.

IoT is a framework used to get the data from the human body's actions and monitoring the continuous activities of them with the help of sensors. Sensors are used directly or indirectly on the human bodies and worked analytics process based on the programming written on that for processing. Notifications will be sent to the person's home or hospital for making immediate treatment to them in a quick manner [10–12]. All details are stored in a cloud server for future analytics and give a detailed report to the society about how this virus will affect and which area, zone, etc. To analyze the vigorousness of the COVID-19 virus, countries have studied various methodologies to provide the solutions but unfortunately, time consumption is very high. They keep trying to manage the world people's data in whole or country wise or day wise with a cloud server which always researching analytics part. Zonal levels are identified in every country namely normal, affected, and heavily affected [4, 13, 14]. They framed rules and precautions for the zones separately to quarantine them easily in their living areas. Finally, there is two living life scenario such as pre-pandemic and post-pandemic which provides the overall view/ idea about the harmful crucial virus effects [15]. In this pandemic period, all the peoples of the countries should strictly follow the precautions said by the scientists and doctors then avoid unwanted gathering in crowded places [16]. COVID-19 virus would create a vital impact on people's day-to-day life and practiced them to live a healthy life by following the proper precautions and treatments.

3 Proposed Framework Architecture

The proposed architecture contains sensors on the IoT framework such as temperature, blood pressure, blood oxygen saturation, and ECG. All the sensors are combined as a mobile holding device unit or a wristwatch model for easy wearing in the human body. The power will be given to the entire circuit by either a chargeable unit or batteries. An Arduino microcontroller is used to control all the sensors, and connectivity establishment has done with Bluetooth or Wi-Fi module inbuilt on that. GSM and GPS modules are attached to this circuit for sending notifications as a location and alert messages to the users. All data generated from human bodies are stored in a local software database and then sent to the cloud server for analytics. Cloud databases are used to produce the threshold values of all symptoms of COVID-19 such as fever, oxygen level, heartbeat level, and blood pressure also. Based on these values, a person who is roaming a lot of places can self-tested with wearable device support and sent notifications to their home. The number of devices is connected to a body will not affect human nature because all are wearable and low-frequency used sensors. Since it is wearable, it can be carried everywhere with protected mode. Figure 2 explains the proposed architecture and its components.

There are several sensor units used in this proposed system and will give exact values from the human bodies for analytics purpose. Table 1 denotes the equipment used in this proposed system and their technical specifications along with justifications.

3.1 Cloud Data Repository

The cloud data base repository has been created with Amazon EC2 tools with minimum space required to collect the information from the wearable devices on spot. Connectivity between the device and cloud was made using Wi-Fi module, and

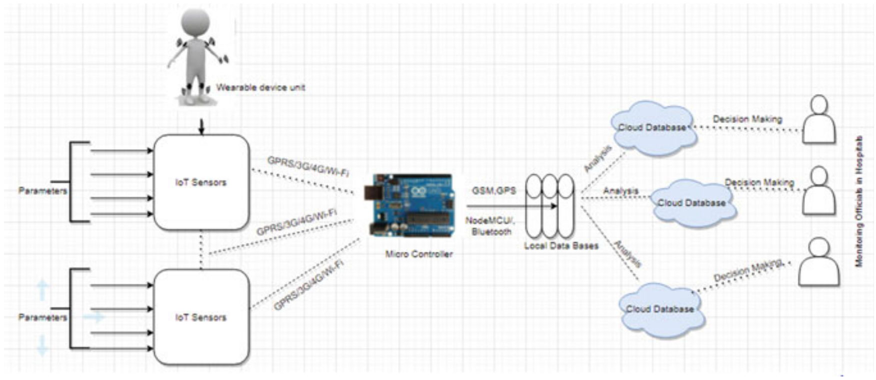


Fig. 2 Proposed architecture

Table 1 Units description used in proposed system

S. No.	Equipment	Details	Justification
1	Arduino	ATmega2560 microcontroller	To control the entire system
2	Temperature sensor	CJMCU-4466 MAX4466	To detect the temperature
3	Mobile Holder	–	To detect the presence of mobile
4	Nodemcu Module	ESB 8266	Sends messages using IOT
5	GPS Module	NEO6M	For location details tracking
6	Bluetooth module	12 V	To control the car ignition switch
7	GSM module	SIM900A	To send an alert message
8	ECG Sensor	ad8232-ecg-sensor	Take heart beat values
9	Blood Pressure sensor	BMP180 Barometric Pressure Sensor	To find the pressure level
10	Blood Oxygen saturation Sensor	Contac CMS50DL	To find the oxygen level

the decision has been made after the analytics part done at data centers of cloud repository. The cloud analytics part has done using machine learning techniques in order to identify the symptoms specified in the training data sets already stored. Based on the threshold value set in the repository trained data, the result will give the decisions about patient's health. Machine learning techniques are used to predict the COVID-19-affected persons from the trained data sets which are already collected and stored in a data repository for analytics. Multiple factors and parameters are taken for decision-making about the patients' health condition periodically collected from different places and persons through the IT wearable devices.

Monitoring officials are available in the hospitals and the places where these devices are using. Their responsibilities are identifying the most affected persons from the different done and contact them using their GPS locations during analytics period. GSM unit is used to send the notifications to these officials about those who are affected by COVID-19 virus. The collection of data has matched with the previous data stored as training data in the repository was done by ML techniques used in Cloud repository. The data model is refined by the deep learning techniques periodically to avoid the conflicts from the device data. A dashboard is used in local database maintain software, and it shows all the detailed analytics and classification about the affected person details.

3.2 Prediction of Cases Analysis

Total of 150 samples are stored in a repository as a training set, and the inputs generated from various persons on different locations can be captured then compared with these sample training sets with threshold values. It is also called as instances that are stored in a repository to make decisions about the patients' health condition immediately. Machine learning algorithms are used to predict the symptoms and their matched threshold values through the wearable devices. That part will be explained in experimental section with their threshold values. Simulation of all the values is collected from device and matched with the already stored data in the cloud repository.

3.3 Assessment Based on Performance

Assessment can be done by the factors called root mean square error, Receiver Operational Characteristics (ROC) area, accuracy, and the estimated time. For computing the data, cross validation and uncertainty matrix are the techniques used in the cloud repository.

3.3.1 Confusion Matrix Representation

A 2 * 2 matrix can be created with computed class and real class as a row and column for validation. Four possibilities as values considered for calculation.

- Positive True Value (PT)
- Positive False Value (PF)

- Negative False Value (NF)
- Negative True Value (NT)

PT has declared as appositive values and mentioned as positive in a design model then PF has decided as negative but it has labeled as positive instances initially. NF can be denoted as negative values instances that are defined positives, and NT has declared as both negative instances.

3.3.2 Calculation of Cross-Validation

It is a mathematical approach used to classify the efficiency of learning from the inputs in the repository as instances. One instance can be used and it will be broken into 10 instances for verification but each instance is doing 15 times iterations for getting accurate result. Every time, the same method can be executed in both computed class and real class members and output is taken for analytics for the decision-making purpose.

Accuracy of the each iteration about the instances is calculated by the Eq. 1 where all the values are taken from the different instances.

$$\text{Accuracy} = \frac{\text{PT} + \text{NT}}{\text{PT} + \text{NT} + \text{PF} + \text{NF}} \tag{1}$$

Root Mean Square Error (RMSE) can be calculated with the average values of all instances considered from the iterations. That value will be calculated with Eq. 2.

$$\text{RMSE} = \frac{\sqrt{\text{PF} + \text{NF}}}{\sqrt{\text{PT} + \text{NT} +}\,\text{PF} + \text{NF}} \tag{2}$$

Receiver Operational Characteristics (ROC) can be calculated using the classifiers that are used for matching the instances calculated in the matrix. The classifier accuracy was calculated against the positive false rate and the positive true rate. It can express by Formulas 3 and 4.

$$\text{Positive True} \; = \frac{\text{PT}}{\text{PT} + \text{NF}} \tag{3}$$

and

$$\text{Positive False Rate} \; = \frac{\text{PF}}{\text{PF} + \text{NT}} \tag{4}$$

4 Work Flow

This system is entirely used to predict the early stage COVID-19 virus-affected persons in their living areas based on the symptoms declared with threshold values of all. A person traveling from one place to another place wearing the IoT framework model device, that person's BP, heartbeat, oxygen level, and temperature levels are measured by the sensors inbuilt in this device. Then, it is stored and accessed by the local software installed on this device. If it matches the already fixed threshold value, then the person should not go anywhere and they will be considered as an affected person. If not their details will be stored in the database for further processing, they have to use sanitizer and wearing masks continuously to avoid spread from others. If the person traveling to a long distance and not maintained the precautions suggested by this system, then the oxygen and temperature level of that person will be decreased in various zones. The location details are tracked by the GPS and if they will affect, then the notification will be sent to the hospital or home immediately. The important instruction given to the persons is to keep wearing a mask, sanitizing their hands frequently, and maintain social distancing between the peoples in crowded places. Figure 3 summarizes the work flow of the entire proposed system in a detailed manner.

5 Experimental on Zone Level Setup

The experimental setup is taken with several persons wearing the devices and asks them to travel to different places for tracking. Normally, a person wearing a mask and sanitizing their hands frequently will never affect. But if they are not maintaining social distance, then the affected level is increased. Based on this, the zones are classified into three levels namely normal zone, affected zone, and heavily affected zones. The last zone is a danger zone, and in that, all the peoples are affected be in quarantine. The personal details are stored in a local database which is installed on inbuilt software for threshold level checking and sent to the cloud server for further analytics. The distance between the peoples is 6 m range will not be affected whereas less than 3 m will be in the affected zone. This will give notification to the hospitals about the person's affected rate, and it will be helpful for them to take immediate treatment. Figure 4 describes the detailed view of zone levels.

Various symptoms have collected from the persons on their zones, and it will be calculated immediately with the local software database and given a detailed treatment procedure or pre cautionary tips to the persons on that spot. Tables 2 and 3 [17–19] summarize the symptoms and its threshold levels for the COVID-19 patient's gender wise.

From the above tables, the affected rate of COVID-19 in different zones resembles a bad impact on male peoples heavily due to roaming and not maintaining the precautionary systems. Female persons are affected low in these zones due to the healthy

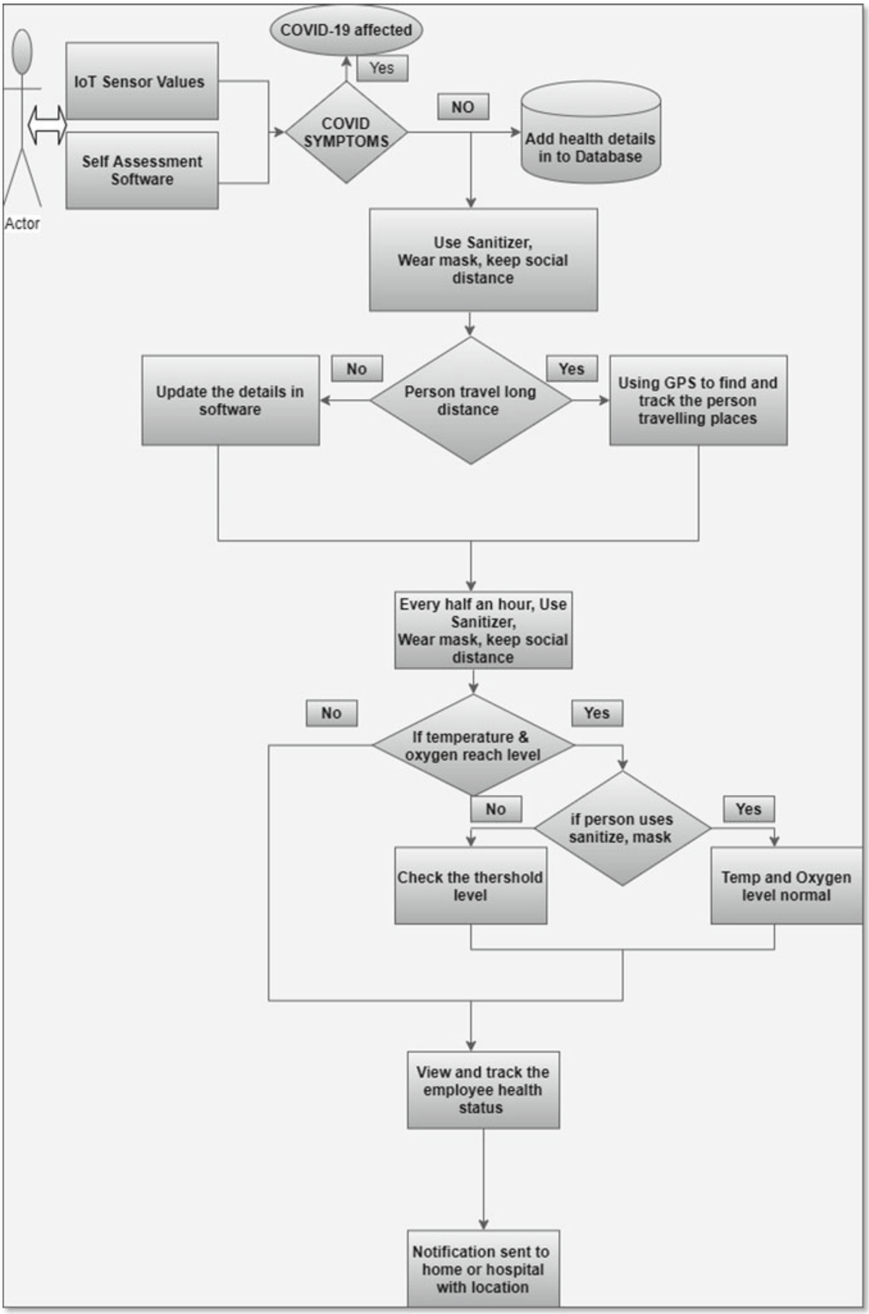


Fig. 3 Work flow of proposed architecture

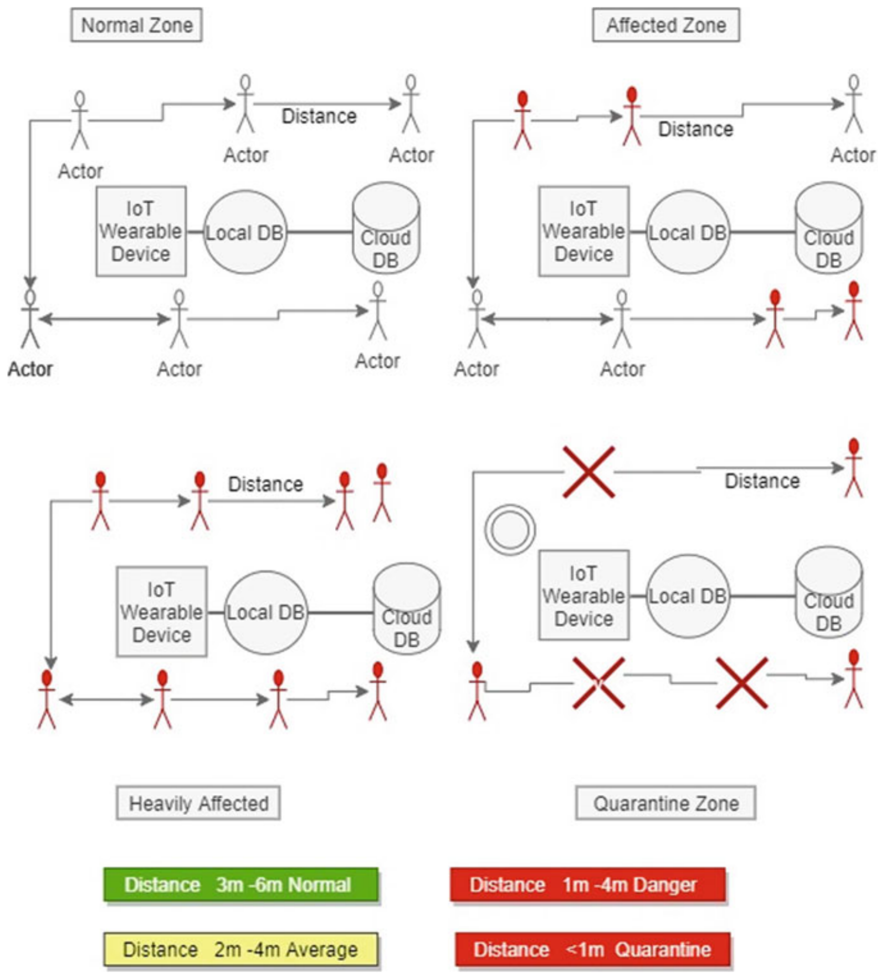


Fig. 4 Zone level person tracking system

foods for increasing immunity powers. We have not considering this test for children and aged persons because the sensors available in wearable devices directly dealing with bodies may cause some side effects. We would like to give suggestions to the peoples that they are not coming out of the home regularly if they come wearing a mask, sanitizing hand, and maintaining social distances are the primary precautions taken for avoiding the spread of viruses. The threshold level varies for children and aged peoples but the precautionary procedures are common for all aged peoples in a country.

Table 2 Symptoms threshold levels for male

Description	Non-COVID	COVID	Threshold level	Affected rate	Actions taken if threshold level reached
Gender of the person	Male	Male	–	High	Avoid roaming
Age of the person	18–45	18–45	> 18 < = 45	High	Take care of Stress, Sugar, BP
Temperature level	92–98.6 F	>100 F	98.6–99.6 F	High	Take test for COVID-19
Saturation level of oxygen	95–100	<90	94	High	To be hospitalized
Number of beats	72–76	70–78	<70	High	High BP, and to be hospitalized
Pulse rate level	72–80	90–120	60–80	High	Need oxygen
Cough level (5 s)	2–3	7–12	>8	High	Take treatment
Blood Pressure	120/80	150/95	130/85	High	To be hospitalized

Table 3 Symptoms threshold levels for female

Description	Non-COVID	COVID	Threshold level	Affected rate	Reason for affected/Not affected
Gender of the person	Female	Female	–	Low	Improve immunity power
Age of the person	18–45	18–45	>18 < = 45	High	Take care of Stress, Sugar, BP
Temperature level	92–98.6 F	>100 F	98.6–99.6 F	High	Take test for COVID-19
Saturation level of oxygen %	95–100	<90	94	High	To be hospitalized
Number of beats	72–76	70–78	<70	High	High BP, and to be hospitalized
Breathing shortness (Blood Pressure)	72–80	90–120	60–80	High	Need oxygen
Cough level (5 s)	2–3	7–12	>8	High	Take treatment
Blood Pressure	120/80	150/95	130/85	High	To be hospitalized

6 Results and Discussions

The experimental setup results taken from the different zones are visualized using modern tools and based on those precautionary methods will be suggested to the affected peoples. One month of data has taken from the person datewise, where they have gone, and how long they have traveled all are tracked with the help of GPS systems and will be stored in a database locally with customized software. Finally, the analytics made like how many cases are active, positive, cured, and death from these results. Since the data has been taken in real time we cannot predict it at the time itself. We have to do analytics then only it is possible to find the solutions. On time, it may measure the threshold level of all symptoms values and informed the hospitals if they crossed.

6.1 *Calculation of Confusion Matrix*

The values taken from the 10 instances with 10 classifiers in one iteration can be calculated and put it in matrix of 2 * 2 format. Positive values are represented higher values, and it is represented in upper-left and lower-right boxes. Low values are represented in lower-left and upper-height boxes for easy identification. Figure 5 denotes all.

635	2
106	753

(a)

629	4
108	759

(b)

639	37
110	748

(c)

528	3
108	719

(d)

527	2
109	810

(e)

653	3
109	826

(f)

605	80
103	829

(g)

598	85
108	853

(h)

518	79
258	865

(i)

628	405
59	859

(j)

Fig. 5 **a–j** Confusion matrix values for different instances

6.2 Performance Measurement Value

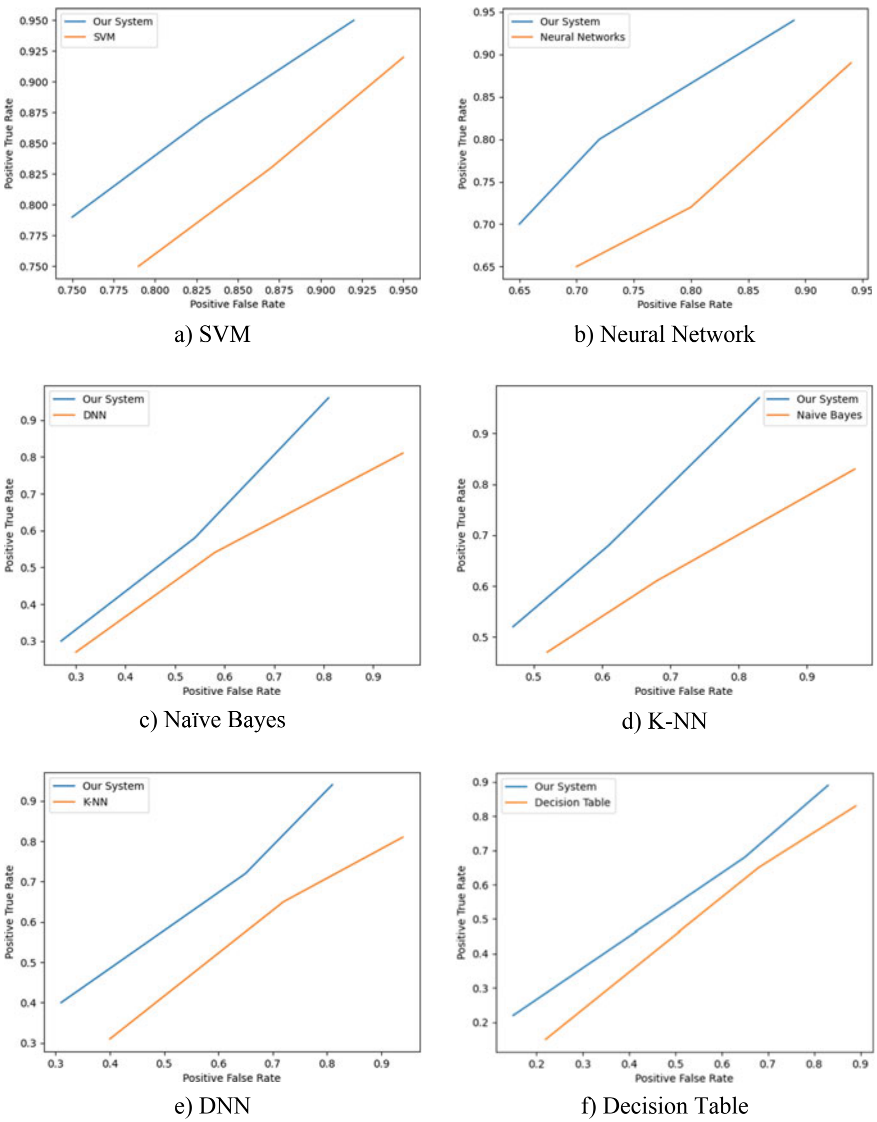
Every value is measured from the 10 classifiers and 10 instances through the cloud repository values, and it is performing the cross validation function with the samples. Each are given the values of RMSE, ROC fields as curve, and their accuracy levels tested by various popular algorithms known as SVM, Naïve Bayes, Neural Network, K-Nearest Neighbor, and Decision table algorithms with our new approach. Out suggested IoT-based system is compared with all these algorithms, and results were noted as curve in Fig. 6a–f.

The above figures represent the ROC curves for the classifiers used in different techniques like SVM, Neural Network, Naïve Bayes, KNN, DNN, and decision table with the new approach IoT wearable device system. Figure 6a–f denotes the difference of the ROC when the positive false rate and positive true rate calculated form the IoT framework and stored in a cloud repository. In each figure, the new approach used by IoT framework has given more positive rates for identifying the right person who has affected the COVID-19 virus using confusion matrix. The normal symptoms and high level symptoms can be classified in every technique, so that the decision-making system is very quick when compared with the previous techniques. Here, machine learning techniques are used to check the modifications in the input data whereas pre-defined training data sets are compared with new data.

Figure 7 describes the performance measures of the entire system build with IoT. There are lot of criteria taken for consideration but the factors such as accuracy, RMSE, and ROC was considered. More or less all the techniques are given the nearest values of the accuracy but the new approach has given more due to the IoT system designed for handling that positions. The other factors like RMSE have given poor values initially but later new trends had come and their capacity has increased in performance. Similarly ROC value has been compared with all old techniques with our new system, for getting better results from the IoT environment devices.

7 Conclusion

The COVID-19 virus spread can be avoided by taking precautionary methods suggested by the scientist but maintained the procedures continuously is a challenging task for all the peoples. Moreover, suggestions are given to a common society of peoples for fighting against pandemic life. This research article provides the methodology to maintain the data which are generated by the human bodies of normal peoples and taken for analytics purposes with the threshold values defined by the standards. When tracking of persons' day-to-day life is monitored regularly, it will be doing good social impact on the peoples who are aware of this pandemic situation. This work implemented the patient's parameters of affected factors and keeps monitoring to avoid the effects. A lot of new innovative methods have come



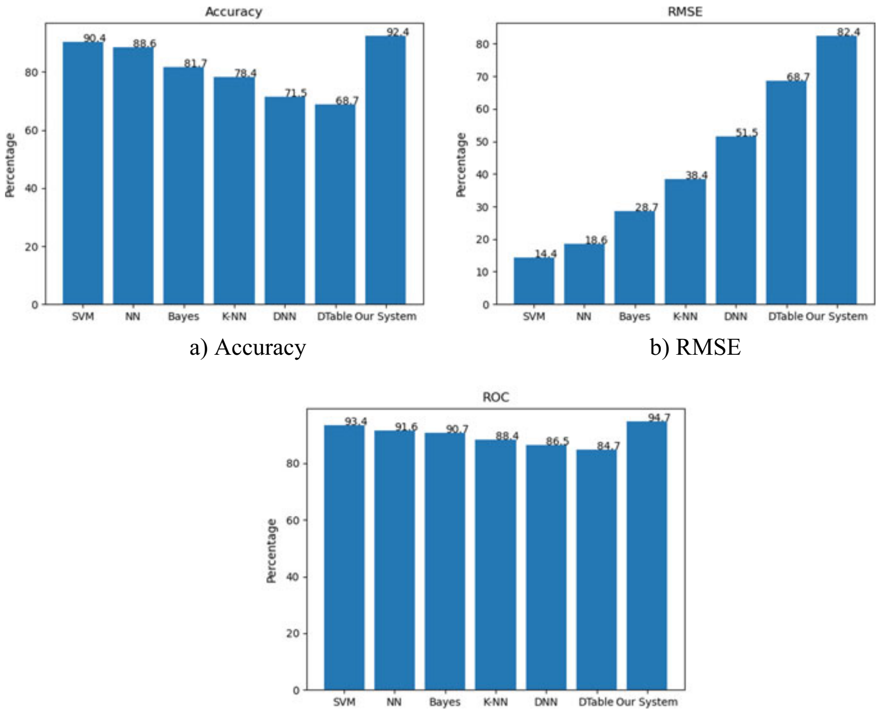
a) SVM b) Neural Network

c) Naïve Bayes d) K-NN

e) DNN f) Decision Table

Fig. 6 a–f ROC comparison

across this pandemic situation but this method provides a better to detect early prediction of COVID-19 with IoT wearable devices. It will give a cost-effective system and carried to all places easily. Maintenance of this device is easy to handle, and power to this system is either rechargeable or battery backup generally. This system provides effective data management framework in IoT during this pre-pandemic life. This system categorized the affected people's day-to-day actions and advised them not to



a) Accuracy

b) RMSE

c) ROC

Fig. 7 **a**–**c** Performance measures

enter into various zones often. The entire system has given the high rate of detection of virus-affected persons and response to the users very less time. Notifications are sent to a limited number of persons with the location details immediately.

8 Future Enhancement

The proposed system in this research work is handled only with limited data since the person is traveling to limited areas and distances. In the future, this system will be working on longer distances and a huge volume of data generated by a single person. Moreover, this system is not advisable for children and aged persons due to the number of sensors used in wearable devices. Next level, we have to plan and design for them also to easily wear this device. Generally monitoring and controlling used in this system by GPS, GSM modules attached the device. The future device consists of several more sensors to predict the disease in early stage not only the COVID-19 virus. The notifications sent by this system are limited in the number of peoples responding. In the future, we have to increase more number of persons to contact, especially the nearest persons who will come and give first aid immediately to avoid the delay. This application-oriented research work is further carried out for blind people also to include various cloud storage modules [20, 21] used for their convenience. Modern tools are used to visualize the predicted data from wearable devices and write suitable algorithms to extract the information from them with high speed and accurate values. The bid data analytics technique device will be implemented in this and used for all kinds of peoples managed data using IoT framework effectively [22].

References

1. Ahmed I, Ahmad A, Jeon G (2020) An IoT based deep learning framework for early assessment of Covid-19. IEEE Internet Things J
2. Lewnard JA, Lo NC (2020) Scientific and ethical basis for social-distancing interventions against COVID-19. Lancet Infect Dis 20(6):631–633
3. Setti L, Passarini F, De Gennaro G, Barbieri P, Perrone MG, Borelli M, Palmisani J, Di Gilio A, Piscitelli P, Miani A (2020) Airborne transmission route of COVID-19: why 2 meters/6 feet of inter-personal distance could not be enough
4. Vedaei SS, Fotovvat A, Mohebbian MR, Rahman GM, Wahid KA, Babyn P, Marateb HR, Mansourian M, Sami R (2020) COVID-SAFE: an IoT-based system for automated health monitoring and surveillance in post-pandemic life. IEEE Access 8:188538–188551
5. Roy A, Kumbhar FH, Dhillon HS, Saxena N, Shin SY, Singh S (2020) Efficient monitoring and contact tracing for COVID-19: a smart IoT-based framework. IEEE Internet Things Maga 3(3):17–23
6. Sartini C, Tresoldi M, Scarpellini P, Tettamanti A, Carcò F, Landoni G, Zangrillo A (2020) Respiratory parameters in patients with COVID-19 after using noninvasive ventilation in the prone position outside the intensive care unit. JAMA 323(22):2338–2340
7. Whitelaw S, Mamas MA, Topol E, Van Spall HG (2020) Applications of digital technology in COVID-19 pandemic planning and response. The Lancet Digital Health
8. Bai L, Yang D, Wang X, Tong L, Zhu X, Zhong N, Bai C, Powell CA, Chen R, Zhou J, Song Y (2020) Chinese experts' consensus on the internet of things-aided diagnosis and treatment of coronavirus disease 2019 (COVID-19). Clin eHealth 3:7–15
9. Allam Z, Jones DS (2020, March) On the coronavirus (COVID-19) outbreak and the smart city network: universal data sharing standards coupled with artificial intelligence (AI) to benefit urban health monitoring and management. In: Healthcare, vol 8, No 1, p 46. Multidisciplinary Digital Publishing Institute
10. Abdali TAN, Hassan R, Aman AHM (2021, January) A new feature in Mysejahtera application to monitoring the spread of COVID-19 using fog computing. In: 2021 3rd International cyber resilience conference (CRC). IEEE, pp 1–4
11. Patrão RL, da Silva CS, da Luz GP, de Caldas Filho FL, de Mendonca FL, de Sousa RT (2020, December) Technological solution development during the COVID-19 pandemic: a case study in an IoT Lab. In: 2020 IEEE global conference on artificial intelligence and internet of things (GCAIoT). IEEE, pp 1–6
12. Awais M, Raza M, Singh N, Bashir K, Manzoor U, ul Islam S, Rodrigues JJ (2020) LSTM based emotion detection using physiological signals: IoT framework for healthcare and distance learning in COVID-19. IEEE Internet Things J
13. Lai X, Liu Q, Wei X, Wang W, Zhou G, Han G (2013) A survey of body sensor networks. Sensors 13(5):5406–5447
14. Islam T, Mukhopadhayay SC (2017) Wearable sensors for physiological parameters measurement: physics, characteristics, design and applications. In: Wearable sensors: applications, design and implementation, pp 1–31

15. Sankar S, Ramasubbareddy S, Chen F, Gandomi AH (2020, December) Energy-efficient cluster-based routing protocol in internet of things using swarm intelligence. In: 2020 IEEE symposium series on computational intelligence (SSCI). IEEE, pp 219–224
16. Sennan S, Ramasubbareddy S, Luhach AK, Nayyar A, Qureshi B (2020) CT-RPL: cluster tree based routing protocol to maximize the lifetime of internet of things. Sensors 20(20):5858
17. Sathiyamoorthi V (2016) A novel cache replacement policy for web proxy caching system using web usage mining. Int J Inf Technol Web Eng (IJITWE) 11(2):1–13
18. Sundarakumar MR, Ganesan N (2016) Real time traffic management system using cloud infrastructure with edge clustering mechanism. Int Res J Eng Technol (IRJET) 3(11)
19. Sundarakumar MR, Mahadevan G (2019) Authorization for secured cloud storage through SHA-256. Int Res J Eng Technol 3:1294–1298
20. Sundarakumar MR, Mahadevan G (2019, September) Improving speed and accuracy of image retrieval using elastic search and features nearest neighbor search. Int J Recent Technol Eng 8:905–909
21. Kumar S (2020) Relocating big data among information centers over the online cloud environments. Available at SSRN 3561400
22. Kumar MS (2021) Design and development of automatic robotic system for vertical hydroponic farming using IoT and big data analysis. Turkish J Comput Math Educ (TURCOMAT) 12(11):1597–1607

Computational Complexity and Analysis of Supervised Machine Learning Algorithms

Jarnail Singh

Abstract Data is generated at a much faster pace, and it is increasing exponentially day by day. Machine learning methods are being used to extract patterns and trends from data to streamline different business activities for more profit with fewer resources. Machine learning models need to be trained with lots of data before being deployed for predictive analysis (Lecture notes in Computer Science, 2012 [1]). Training time depends upon the complexity of an algorithm. We are analyzing the space and time complexity of various machine learning algorithms so that it becomes easier to select and deploy the most efficient and appropriate model for a particular dataset. This research work primarily focuses on data analytics for supervised machine learning algorithms in industrial research domains.

Keywords Time and space complexity · Linear regression · Logistic regression · SVM · Decision tree · KNN

1 Introduction

Machine learning experts often find it difficult to select the most suitable algorithm for the given problem and dataset. Time and space complexity plays a very crucial role in selecting the most efficient machine learning algorithm [2]. The space complexity of an algorithm is the total memory space required by the algorithm for its execution, and it is stated as a function of input size [2]. In simple words, memory space is required by the algorithm to execute from start to finish the task. The time complexity is a measure of the number of operations performed from the beginning of the algorithm to completion of the job and is also expressed as a function of input size. In other words, the time required by an algorithm to complete the task. Machine learning algorithms are used for extracting useful trends and patterns from the vast amount of data that could be used for predicting unseen data. Hence, machine learning algorithms

J. Singh (✉)
University Institute of Computing, Chandigarh University, Mohali, Punjab, India
e-mail: jarnail.e9937@cumail.in

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_16

are nowadays being deployed everywhere to streamline business activities by optimizing the resources being used. Many big organizations are using machine learning algorithms for classification, clustering, image recognition, object detection, image segmentation, computer vision, natural language translation, etc. These machine learning algorithms are trained using huge dataset for capturing the underlying logic of the dataset. Training the machine learning model consumes a substantial amount of computing resources such as time and space [3, 4].

The runtime of an algorithm is not only crucial to comprehend the inner workings of the algorithm but also yields a more successful implementation. Therefore, it is necessary to analyze various machine learning algorithms so that effective and efficient algorithms can be selected for various problems in different domains having varied sizes of the dataset.

The paper is organized into two sections. Sect. 2 is investigating training time and testing time complexities of various supervised machine learning algorithms.

Section 3 investigates about the suitability of major machine learning models based on the size of the dataset, accuracy of model, linearity, execution speed versus training time, and multiple features.

2 Complexity of Machine Learning Algorithms

In this study, the training and testing time complexity of machine learning algorithms is investigated with respect to the size of dataset, number of features, number of nearest neighbors, etc. [1].

2.1 K-nearest Neighbors (KNN)

It is assumed in the KNN algorithm that similar things exist on the brink of one another. In other words, similar things have proximity to every other.

There is no explicit training phase for KNN; hence, KNN does not learn any function that maps input data to output label during the training stage [2]. It just calculates Euclidean/Manhattan distances and sorts the data accordingly. Due to this, KNN is also called a slow learner. At the time of classification stage, when new observation (*xq*) is given, then it classifies the new data into a category that is much similar to it (Fig. 1).

There are two main techniques in KNN:

(1) Brute force KNN or regular KNN and (2) *k-d* tree KNN.

Where n is the count of observations in the dataset, d is the number of attributes, and k is the count of nearest neighbors.

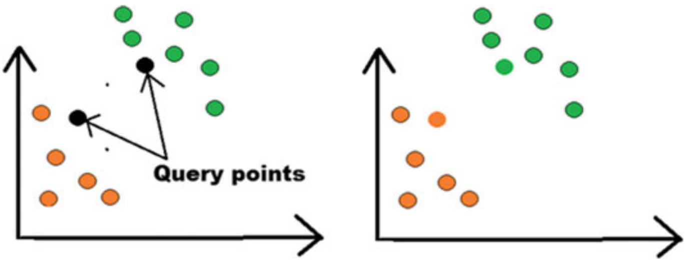


Fig. 1 KNN (*source* https://www.analyticsvidhya.com)

(1) **Brute force KNN**:

The training time and training space complexity are $O(knd)$ and $O(nd)$, respectively, whereas testing time and testing space complexity are $O(knd)$ and $O(n\mathrm{d})$, respectively [2].

(2) **k-d tree KNN**:

The training time and training space complexity are $O(nd)$ log(n), respectively, whereas testing time and testing space complexity are $O(k \log(n)\, d)$ and $O(nd)$, respectively [2].

2.2 Naive Bayes

It is a probabilistic classifier that assumes that there is a strong independence between features [3].

Suppose, we are given k different classes, namely $CK = (c1, c2, \ldots, Ck)$ and 'n' different observations as $x1, x2, x3, \ldots, xn$.

The model is stated as *un* per

$$\begin{aligned} p(Ck \mid x1, \ldots, xn)\, \alpha\, p(Ck, x1, \ldots, xn) \\ \alpha\, p(Ck)\, p(x1 \mid Ck)\, p(x2 \mid Ck)\, p(x3 \mid Ck) \\ \alpha\, p(C_k) \prod_{i=1}^{n} p(xi \mid Ck) \end{aligned} \tag{1}$$

During the training stage, the likelihood and class probability are calculated.

During the testing stage given a query point xq, the class having the maximum probability of the query point is selected, and the query point (xq) is classified into the class having the maximum probability.

We assume that the number of observations in the dataset is n, number of attributes is d, and C is the number of classes.

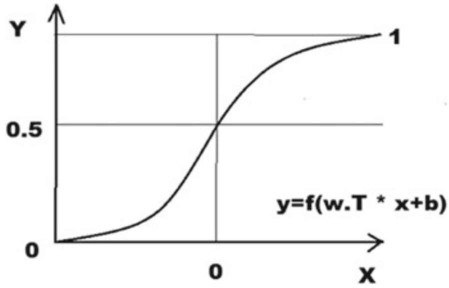


Fig. 2 Sigmoid function (*source* https://www.analyticsvidhya.com)

The training time and training space complexity of naïve Bayes classifier are $O(ndc)$ and $O(dc)$, respectively. Testing time and testing space complexity are $O(dc)$ and $O(dc)$, respectively [2].

2.3 Logistic Regression

Logistic regression classification model assumes that different classes are separable by straight line very clearly.

The sigmoid function is used in this model for classification.

$$f(x) = 1/(1 + e^{(-x)}) \tag{2}$$

When we plot the graph of the sigmoid function, then it will be as under (Fig. 2).

Where sigmoid function is denoted by $f(x)$. The sigmoid function value is bounded by 0–1. The sigmoid function is mostly used as it is not prone to outliers, and it can have a very nice probabilistic interpretation [2].

The discriminative function is learned in the training phase. During the training phase, the associated weights with each feature can be learned. If xq be our query point in the test stage, then [5]

$$Y = f(w.T * xq + b) \tag{3}$$

Both the training time and training space complexity of the logistic regression are $O(nd)$, whereas both the test time and test space complexity of logistic regression are $O(d)$ [6–8].

2.4 Linear Regressions

It assumes that there is a linear mapping between the independent input variable and dependent output variable. It learns the function that maps input to output label

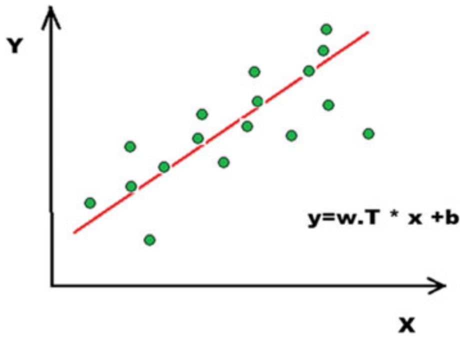


Fig. 3 Linear regression (*Source* https://www.analyticsvidhya.com)

during the training stage [2, 3]. In the training phase, we learn regression coefficient (w) and regression constant by iterating through each observation (Fig. 3).

Both the training time and training space complexity of linear regression are $O(nd)$, whereas both testing time and testing space complexity are $O(d)$ [2, 3].

2.5 Support Vector Machine (SVM)

It assumes that different classes are clearly separable with straight lines (Fig. 4).

The margin of the hyperplane is maximized so that new data points could be clearly classified into a particular class [4]. The alpha (α) is learned with respect to each observation and then use the observations which have the value of alpha greater than 0 as support vector. Hence, only observations that are important are the support vectors, and due to this, it is named as support vector machine.

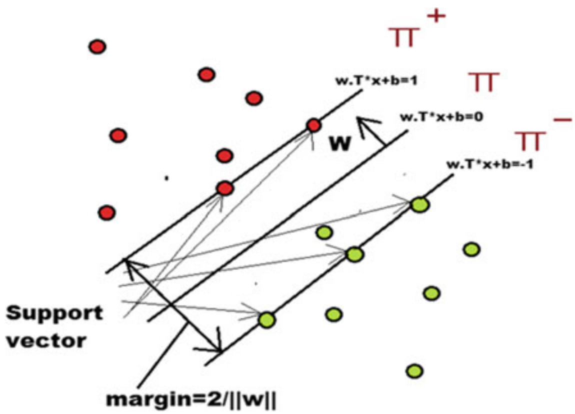


Fig. 4 Support vector machine (https://www.analyticsvidhya.com)

We use the kernel trick if data points are not linearly separable. The feature conversion is where we transform observations indirectly to higher dimensions, and thereafter data points could be separated linearly.

The value of support vector and their corresponding alpha values are calculated during training phase. Suppose *xq* be the new query point during test time.

$$f(x) = w.T * xq + b$$

$$w = \sum_{n=1}^{n} \alpha iyixi$$

$$f(x) \sum_{n=1}^{n} \alpha iyixi * xq + b$$

If count of support vectors is *k,* then training time and training space complexity are $O(n^2d^2)$ and $O(nd)$, respectively, whereas both testing time and testing space complexity are $O(kd)$ [6].

2.6 Decision Tree

The decision tree is built while training. The feature with the highest information gain or highest entropy reducing feature or with the highest Gini impurity is selected.

The training time and training space complexity are $O(n \log nd)$ and O(node count), respectively, whereas test or runtime and test or run space complexity are $O(\log n)$ and O(count), respectively [7].

If there are *n* leaf nodes in the decision tree, then depth of tree is log *n,* and the total number of nodes is $2 * n - 1$.

If the number of leaf nodes is 8 (Fig. 5)

If the number of leaf nodes is 4 (Fig. 6)

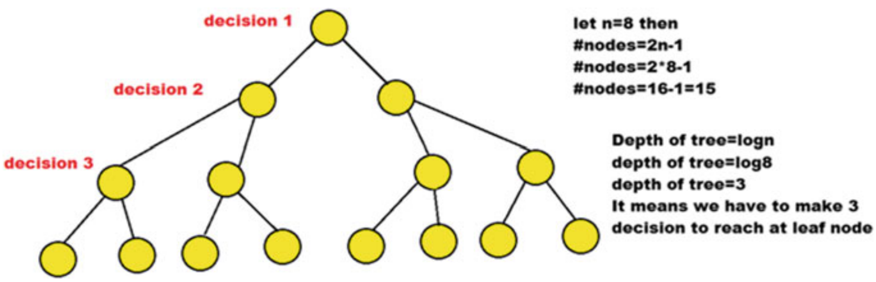


Fig. 5 Decision tree (*source* https://www.analyticsvidhya.com)

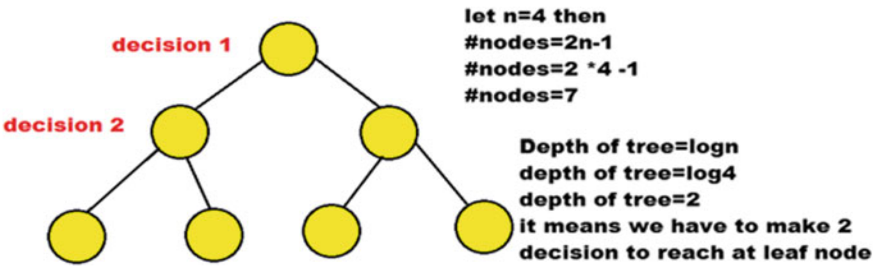


Fig. 6 Decision tree (*source* https://www.analyticsvidhya.com)

2.7 Ensemble Models

The ensemble model in machine learning is to combine the multiple models to enhance the prediction power of the model.

(1) Bagging is an iterative process of reducing the error by deploying multiple models in sequence, which is called random forest. One of the best examples of bagging is random forest. There are two popular techniques supported by bagging, bootstrapping and boosting.

In random forest, low bias and variance are required for decision trees, which will make our trees to be overfit the training dataset. The decision tree that is either full or has high depth will have high variance.

If there is K number of trees in the random forest, then training time and training space complexity will be $O(k\ n \log nd)$ and $O(\#\text{nodes}\ k)$, respectively, whereas testing/running time or testing/running space complexity is $O(k \log n)$ and $O(\#\text{nodes}\ k)$, respectively [8].

(2) **Boosting: Gradient boosted decision tree (GBDT)**

They are the most popular methods in the regression and classification categories for prediction.

Let M be number of tress and Gamma m be output values for each leaf of decision tree, then the training time and training space complexity will be $O(M\ n \log nd)$ and $O(\#\text{nodes}\ M+\text{gamma}\ m)$, respectively, whereas testing time and testing space complexity are $O(M \log n)$ and $O(\#\ \text{nodes} * M + \text{gamma}\ m)$, respectively [9].

Gradient boosted decision tree is less efficient in comparison to the random forest.

3 Analysis

Here, we are considering upper bound complexity for different supervised machine learning algorithms and also implementing them using SKlearn. It is assumed that the complexities take the form $f\ O(n^a p^b)$ and a and b are estimated using randomly generated samples with n (number of samples) and d (number of features) varying.

Sr.NO	Algorithm	Classification /Regression	Training time	Testing time	a	B
1	KNN	C+R	O(nd)	O(k log(n) d)	-	-
2	Naïve Bayes	C	O(ndc)	O(dc)	-	-
3	Logistic Regression	C	O(nd)	O(d)	0.92	0.87
4	Linear Regression	R	O(nd)	O(d)	0.71	1.25
5	SVM	C+R	$O(n^2d^2)$	O(kd)	2.06	0.52
6	Decision tree	C	N log (nd)	O(log n)	0.72	0.53
7	Random Forest	R	O(k n log(nd))	O(K log n)	1.22	0.9

Fig. 7 Comparison of different algorithms using SKlearn for implementation

Then, to compare and estimate the complexity of different algorithms, a log–log regression is used (Fig. 7).

3.1 Training Dataset Size

The reliability of prediction mainly depends upon quantity and quality of dataset. However, in many situations, size of the dataset may be small or the number of features in the dataset is also very few [10]. So, if the training size of the dataset is small as well as if the dataset has more features, then one needs to select algorithms with high bias/low variance such as rectilinear regression, naïve Bayes, or linear SVM for better accuracy. Linear regression is better in terms of training time and testing complexity as compared to naïve Bayes and linear SVM.

If the training dataset has large number of observations as compared to the number of features, then algorithms such KNN, decision trees, or kernel SVM can be opted as these algorithms have low bias and high variance [11]. Decision tree is better in terms of training time as well as testing time complexity as compared to KNN and kernel SVM.

3.2 Accuracy and/or Interpretability of the Output

The model accuracy indicates how closely the mapping function is extracted by the model using the training dataset. Mapping function is the function that transforms the input attributes to output label. A rigid model is less accurate but easier to understand and interpret. For example, linear regression, where one can understand the mapping between predictor attribute to predicted attribute. The flexible models are more accurate but difficult to understand and interpret [9].

Rigid algorithms produce a narrow range of shapes for mapping from input to output. Linear regression is the best example of a rigid approach as it produces only a linear mapping from input to output.

On the other hand, the algorithms that generate functions with a broader range of possible shapes are called flexible. KNN with $k = 1$ is the best example that generates an input to output mapping function which is highly flexible and more accurate (Fig. 8).

Hence, there is a trade-off between simplicity and accuracy. If one wants interpretable models, then one can use restrictive models as they are simpler and more interpretable. On the other hand, if one wants more accuracy, then one can go for more flexible models as they provide better accuracy. Linear regression and decision trees are better options in terms of training time and testing time complexity as well as interpretability.

Fig. 8 Representation of trade-off between flexibility and interpretability using different statistical learning methods. *Source* [12]

3.3 Speed Versus Training Time

The more complex function gives better accuracy, but it will take a longer training time [3]. When the training dataset is very large, then training time increases substantially. Therefore, these two factors affect the selection of algorithms in real-world applications irrespective of domain.

There are simpler, easily interpretable algorithms such as naïve Bayes and linear and logistic regression. The logistic regression has better time and space complexity for training as well as testing. On the other hand, there are algorithms such as SVM that require parameter tuning, neural networks that take longer to converge and random forests that require a longer time to train and test the model. Logistic regression performs better in terms of time complexity of training and testing as well as more efficient.

3.4 Linearity Versus Nonlinearity

Many simpler algorithms assume a straight line can separate observation into different classes. Logistic regression and support vector machines are the best examples of such algorithms [5, 13]. Linear regression algorithms are simple that assume that a mapping function is linear. Often, real-world data may not conform to linear function, but they perform well if there is a linear mapping between input data and output.

For nonlinear data, we may need models which are capable of handling multidimensional data and more complicated mapping functions. The most popular examples of models are kernel SVM, random forest, and neural nets. Random forest outplays in terms of time and space complexity of training and testing the model.

We find out the linear relationship using linear or logistic regression or support vector machine and measure the residual errors. If the residual error is higher, then the data is nonlinear and would require complicated model to map the input to output. Linear or logistic regression outplays in terms of time complexity of training and testing the model.

3.5 Multiple Features

The dataset often has a n number of attributes, out of which a few of them may have explanatory power [14, 15]. The relevant features are kept, whereas irrelevant features are removed to make the model computationally simple and easy to interpret, and it also reduces training time substantially.

If the dataset has large attributes and the number of observations is small, then SVM is well suited because it has better time complexity for testing. Various feature selection algorithms such as PCA are used for dimensionality reduction and to select only the relevant features that have maximum explanatory power.

4 Conclusion and Future Scope

1. For a small training dataset with more features, select algorithms with high bias/low variance, e.g., rectilinear regression, naïve Bayes, or linear SVM. Linear regression is better in terms of time complexity of training and testing the model.
2. For a large training dataset with small number of features, select algorithms with low bias and high variance such as KNN, decision trees, or kernel SVM. Decision tree is better in terms of time complexity of training and testing the model.
3. For higher accuracy, select more flexible model such as KNN or SVM, and for better interpretability, opt for rigid model like linear regression as it would be easy to fine parameters. Linear regression and decision trees are better options in terms of training time as well as testing time complexity.
4. With large dataset for better speed, select more efficient algorithms such as logistics regression, and for better accuracy, select more complex algorithms such as neural networks, random forest, and SVM that take longer to converge. Logistic regression is better in terms of time complexity of training and testing.
5. For linear mapping functions, select simple functions such as linear regression, and for nonlinear mapping, select more complex functions such as random forests, kernel, or neural nets.
6. For a dataset with multiple features, select only relevant features that have explanatory power to make model computationally simple and easy to interpret.
7. Most of the machine learning and deep learning algorithms are difficult to interpret and hence difficult to fine-tune for better accuracy. Therefore, various models should be explored with varying dataset size and varying number of features to understand, interpret, and fine-tune hyperparameters for better accuracy and less training time and testing time.

References

1. Lecture notes in Computer Science (2012)
2. https://medium.com/analytics-vidhya/time-complexity-of-ml-models-4ec39fad2770
3. https://www.kdnuggets.com/2020/05/guide-choose-right-machine-learning-algorithm.html
4. https://en.wikipedia.org/wiki/Naive_Bayes_classifier
5. Komarek P (2004) Logistic regression for data mining and high-dimensional classification. Robotics Institute, p 222
6. Big O notation (2015, May 18) In Wikipedia, the free encyclopedia. Retrieved from http://en.wikipedia.org/w/index.php?title=Big_O_notation&oldid=662932976

7. Big O Notation (n.d.) Retrieved from http://web.mit.edu/16.070/www/lecture/big_o.pdf
8. Big O Notation | Interview Cake (n.d.) Retrieved June 22, 2015, from https://www.interviewcake.com/big-o-notation-time-and-space-complexity
9. https://www.analyticsvidhya.com
10. https://levelup.gitconnected.com/train-test-complexity-and-space-complexity-of-logistic-regression-2cb3de762054
11. Murphy KP (2012) Machine learning: a probabilistic perspective. MIT Press
12. Ke G et al (2017) LightGBM: a highly efficient gradient boosting decision tree. Adv Neural Inf Process Syst 30 (2017): 3146–3154 (Appendix: Springer-Author Discount)
13. https://www.thekerneltrip.com/machine/learning/computational-complexity-learning-algorithms/
14. Bache K, Lichman M (2013) UCI machine learning repository. The University of California, School of Information and Computer Science, Irvine, CA
15. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks science 313(5786):504–507
16. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press

An Intelligent Iris Recognition Technique

Salam Muhsin Arnoos, Ali Mohammed Sahan, Alla Hussein Omran Ansaf, and Ali Sami Al-Itbi

Abstract Biometrics are vital in security. Facial recognition, fingerprints, and iris recognition are all examples of computer vision biometrics. Unique authentication based on iris structure is one of the finest approaches for iris identification. This research provides an iris-based biometric identification system combining CNN and Softmax classifier. The system consists of picture augmentation by histogram equalization, image reduction by discrete wavelet transformation (DWT), segmentation by circular Hough transform and canny edge detector, and normalizing by Daugman's rubber-sheet model. Each picture is adjusted before being fed into the DenseNet201 model. The Softmax classifier then sorts the 224 IITD iris classes into 249 CASIA-Iris-Interval classes, 241 UBIRIS.v1 iris classes, and 898 CASIA-Iris-Thousand classes. The performance of our suggested system is determined by the setting of its deep networks and optimizers. In terms of accuracy, it exceeds existing approaches by 99%.

Keywords Deep learning · DenseNet201 · Convolutional neural network (CNN) · Softmax classifiers · Discrete wavelet transformation (DWT)

1 Introduction

For example, biometrics includes fingerprints and retina scans. Biometric systems have the best track record for person authentication. Because biometrics is difficult to steal, it may be used instead of passwords and smart cards. Behavioral traits are often paired with physical traits to create biometric IDs. Physiological qualities of the body include fingerprints, palm veins, DNA, face recognition, iris, and so on [1].

S. M. Arnoos (✉) · A. M. Sahan · A. S. Al-Itbi
Informatics Department, Technical College of Management, Middle Technical University, Baghdad, Iraq
e-mail: mustafa_kut88@yahoo.com

A. H. O. Ansaf
Kut Technical Institute, Middle Technical University, Baghdad, Iraq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_17

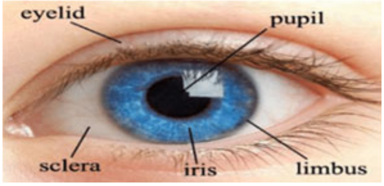


Fig. 1 The outer structure of human iris

Other categories included speaking voice, walking motion, etc. Biometrics helps safeguard people's health in two ways: physically and behaviorally. Biometrics has entered our everyday lives. Although iris is certified for precise biometric systems. Many people believe that biometrics is the proper technique to identify someone. The iris is the annular eye area between the white sclera and the black pupil [2].

Iris texture is very random and distinct, making it unlikely to be unique across two iris patterns or two eyes. Circular, ridged, crypts, freckles, furrows, and zigzag patterns are among the most common iris texture traits. Its reputation for dependability helps in distinguishing individuals. Until his death, the iris pattern remained unchanged. Because recognition techniques are not susceptible to spoofing assaults, they are regarded safer. Again, altering its patterns requires more risk than altering fingerprints, which are simpler to manipulate. Eyelashes, reflections, and eyelids may all interfere with iris-based identification systems. For example, Aadhar card access, banks, and current apps employ iris-based identification [3] (Fig. 1).

This study encompasses the following: Sect. 2, the study of related Studies. In Sect. 3, there is the proposed iris recognition system, and in Sect. 4, there are the experimental results and analysis. Finally, the study's future evolution is summarized in the last section.

2 Review of Related Studies

S. Arora and M. S. Bhatia an iris scan may be used to verify a person's identification using deep learning. The circular Hough transform and CNN is used to locate an iris. Then, using a Softmax classifier, the dataset is classified into one of 224 classifications. The accuracy of their proposed system relies on the choice of hyper-parameters and optimization methods.

To discover the iris border in the eye picture, Reddy [4] used Daugman's rubber-sheet model, followed by a circular Hough transform and edge detection. The recommended system is then trained using a CNN to identify iris characteristics. This classifier uses Softmax to classify 224 iris datasets and 108 CASIA V1 Iris datasets. Thus, the system's overall performance and deep-learning training and optimizations. We met our targets with 98% success and 95.4% accuracy.

Alaslni and Elrefaei [5] suggested CNNs (VGG-16s) for iris recognition model training. They used a pre-trained VGG-16 to successfully train on the IITD and

CASIA Iris datasets. However, the suggested solution does not work with other publicly available datasets or solve other biometric issues.

CNN's iris recognition system was developed by Kien Nguyen et al. They tested DenseNet on two datasets, LG2200 and CASIA-Iris-Thousand, and found that it outperformed AlexNet, VGG, Google Inception, and ResNet. The authors did not explain why the AlexNet model was less accurate.

Shervin Minaee et Deep CNN features for iris recognition: a conceptual study They developed a system using VGG Net characteristics and got decent results. They studied data from CASIA-Iris-1000 and IIT Delhi using two well-known databases. This study uses a modified object recognition framework. Their findings would have been more accurate if they had used an architecture optimized for iris detection.

Aziz Nazri Ali and Sue Chin Yow His idea was to use deep learning to develop an iris scanner. In addition to SVM, Bayesian optimization, and data augmentation, this study uses CNN and data augmentation. The study's dataset is CASIA V1. The results are noteworthy; however, the approaches are not suited for other publicly accessible datasets.

3 The Proposed Iris Recognition System

In Fig. 2, the proposed iris recognition system uses a denseness for feature extraction and classification. The iris recognition technology will be developed in three areas: It includes both picture capture and preprocessing.

3.1 Image Acquisition

After capturing a picture of the iris, it may be preprocessed to improve clarity. CASIA-Iris-1000 [6], interval 12 [7], Delhi 10 [8], and the UBIRIS Database [9] were used in this investigation. The table has not been preprocessed Table 1.

3.2 Preprocessing

The four methods depicted in Fig. 3 are used before processing iris images. Because collected pictures are packed with noise, they must be processed before being fed to segmentation. Because brightness is not uniform and noise must be reduced, a histogram equalization is used to enhance the picture. Use unsharp mask to sharpen your picture to bring out textures and characteristics.

Fig. 2 Flow diagram of the system employed in this study

Table 1 Presents the information on images on each dataset before preprocessing and is known as the information on images before preprocessing

Dataset	Number of classes	Number of samples	Image size in pixel
CASIA-Iris-Interval	249	2639	320×280
CASIA-Iris-Thousand V4	1000	17,938	640×480
UBIRIS.v1	241	1876	200×150
IIT Delhi	224	2240	340×280

3.2.1 Histogram Equalization

Equalization of histograms: A histogram equation is applied to increase the contrast and brightness of images. A linear and equal distribution of colors is obtained by

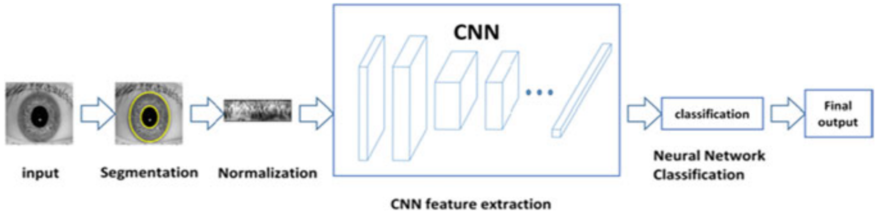


Fig. 3 The process involved in iris systems

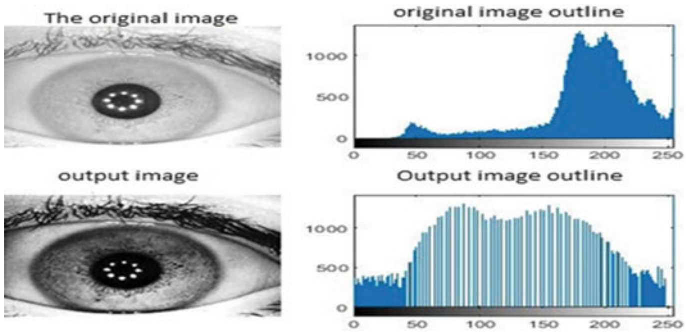


Fig. 4 The histogram equation color distribution

examining the color distribution in the image and choosing a conversion function to modify the light values. The number of repetitions for each color is calculated, and then, the number of repetitions for each color is calculated and done. According to Fig. 4, the values are divided by the sum of the colors in the image to yield new color values.

3.2.2 Unsharp Mask Filter

To highlight the image's texture and detail, the unsharp mask is applied. By improving the resolution of the image at a tiny scale, it can substantially increase the look of detail. It removes an image that has been smoothed out of the original image. You first generate a copy of the original image that is very fuzzy and then you remove the blurry copy from the original image, leaving only the clearer image behind. The remaining difference is known as a mask. Next, you apply the mask to the original image. Completing the original image with the mask, rather than masking the original, leaves behind a clearer end result.

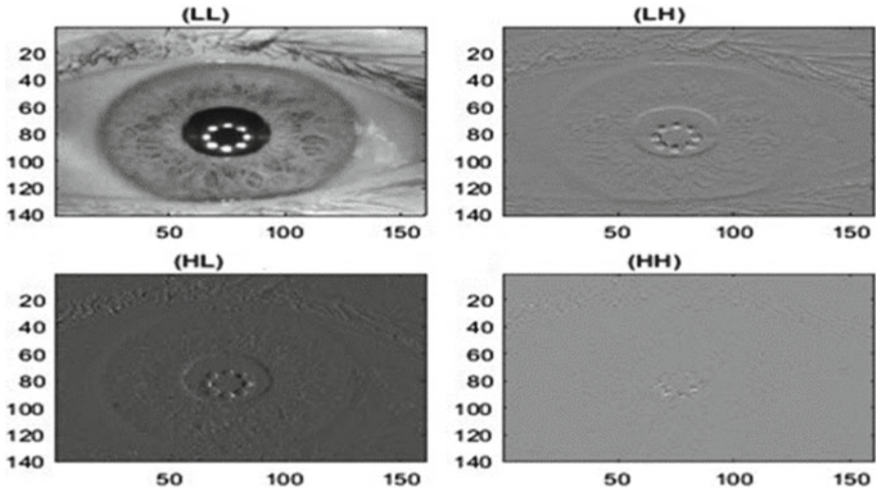


Fig. 5 Iris template using one-level DWT

3.2.3 Reducing Image Dimensions and Preserving Important Features

We can help with preprocessing, a method for reducing picture size and speeding up the matching process. Our first DWT was run on the normalized iris picture. The DWT transformation, shown in Fig. 5, adds four sub-band frequencies to the iris image: LL, LH, HL, and HH. The iris low-frequency coefficients are encoded in the LL sub-band [10], which is used for iris segmentation.

Wavelet discretization: A wavelet transformation reduces noise by decomposing data or images into wavelet coefficients [11]. The inverse discrete wavelet transformations are utilized to reconstruct the picture [12]. Figure 5 shows the resultant frequency sub-bands (LL, LH, HL, and HH). Iris features may be characterized using low-frequency coefficients, LL sub-band, allowing us to employ LL frequency sub-band for feature reduction [10].

3.2.4 Iris Segmentation

The next step in iris identification is iris segmentation, which involves processing a digital eye picture to determine the iris area. Figure 1 illustrates two concentric circles as an iris area approximation. The inner-circle depicts iris/sclera; the outer circle represents iris/pupil. The Hough transformation circular iris area.

The Hough transformation computes geometric parameters (lines and circles) in an image. The Hough transform may be used to identify the iris and pupil center coordinates and radius. Commonly used to identify object shapes based on class names. An edge map is created by calculating the gradients (first derivatives of intensity values) in an eye picture. To obtain the greatest values for circle parameters,

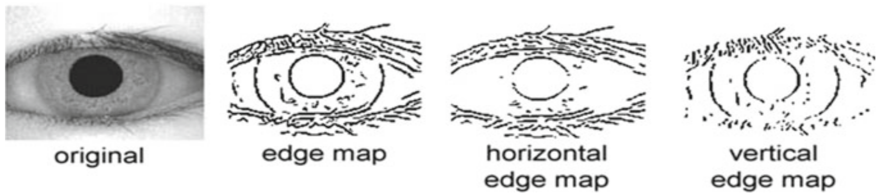


Fig. 6 Segmentation using Hough transformation

neighboring areas on the circle at various radii are evaluated, and votes are cast [13]. This equation computes the center's coordinates and radius:

$$X_C^2 + Y_C^2 - Y^2 = 0 \tag{1}$$

The circle's radius (r) is the biggest number, and the circle's center coordinates (X_C, Y_C) are determined using Hough space edge points. Vertical derivatives (gradients) are used to identify the iris-sclera border. To remove the eyelids' effect from the Hough transform, only vertical gradients may be employed. Some of the circle's boundary pixels are unnecessary. Less edge points to vote on improve accuracy and efficiency [13] (Fig. 6).

3.2.5 Normalization

After successfully segmenting the iris area from an eye picture, the following step is to transform it into a fixed-size rectangular block. Normalization creates an iris with regular proportions. These measures are consistent; therefore, photographs of the same iris will have comparable properties [14]. This work uses Daugman's rubber-sheet model to normalize the iris image.

Rubber-sheet model: As the name implies, the rubber-sheet model transforms the iris into a rectangular block in non-concentric polar coordinates, which is the most prevalent kind [15]. To calculate the comparable location on the polar axis (r), Daugman created a rubber-sheet model. The angular resolution is calculated by dividing the number of radial lines by the number of data points. The radial resolution is the number of data points. Using Eq. 2,

$$I[x(r,\theta), y(r,\theta)] \rightarrow I(r,\theta) \tag{2}$$

(x, y) is the iris position, and (r) is the Cartesian and normalized polar coordinates. $r$ranges from Rp to Rl, while varies from 0 to 2. The boundary points are utilized to calculate $x(r)$ and $y(r)$.

The array's horizontal and vertical dimensions correspond to angular and radial resolution, respectively (Fig. 7).

Fig. 7 Rubber-sheet model of Daugman

3.2.6 Deep Learning for Iris Recognition

After preprocessing, we use a Softmax classifier to extract and classify CNN features. The DenseNet201 model is used in this work to simulate the CNN architecture. This is the top categorization layer; it is linked. The Softmax classifier then gets the fully connected layer's output and computes the probability distribution across *N* classes. This section also discusses the efficient recognition of the training process, network architecture, and input picture size. Keeping over fitness at bay throughout training is vital.

Training Methodology

The database is used to build the training, validation, and testing sets. 80% of the database is for training, and 20% is for testing. The validation set also contains parameters with a low validation error. The Softmax classifier divides the training set into mini-batches and calculates the training errors for each. The validation set is used after each epoch to test the present architecture's validity. The present architecture's cost value and validation error are then determined. This study's beginning learning rate is 0.001. The dropout approach [16] avoids overfitting. Each node and its connection are ignored in each training iteration by 0.2. After the training phase, the best model is chosen based on performance recognition.

Network Architecture

DenseNet is a CNN model created by Huang et al. and won the ILSVRC-2016 competition in 2017. An already trained and configured convolutional neural network model called DenseNet201 is utilized for the iris recognition [17] in this method. Each of the four dense blocks and three transition layers in DenseNet201 is shown in Fig. 8. To enhance the capabilities of the DenseNet201 bottleneck output features, a flattening layer and a dense layer are applied, and then, a Softmax layer is applied on top.

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112×112	7×7 conv, stride 2			
Pooling	56×56	3×3 max pool, stride 2			
Dense Block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv			
	28×28	2×2 average pool, stride 2			
Dense Block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv			
	14×14	2×2 average pool, stride 2			
Dense Block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14×14	1×1 conv			
	7×7	2×2 average pool, stride 2			
Dense Block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1×1	7×7 global average pool			
		1000D fully-connected, softmax			

Fig. 8 Shows DenseNet architecture

Input Image Size

The size of the input image influences the speed and accuracy of the network. To test the system's performance, three input image sizes (128 × 128 pixels, 64 × 64 pixels, and 32 × 32 pixels) are employed. Normalization and histogram equalization followed by three layers of DWT are used to create these photos. As the iris size is lowered, the iris patterns will become completely undetectable; nevertheless, for the larger size, larger memory requirements and the higher level of complexity computations could pose a challenge.

4 Experimental Results and Analysis

It is tested with four iris datasets: IITD Iris Databases [8], UBIRIS.v1 [9], CASIA-Iris-Thousand [6], and CASIA-Iris-Interval [7]. This set of iris photographs uses pupil dilation, eyelid/lash occlusion, a tiny eyelid shadow, and specular reflection. Table 2 lists the dataset's specifications.

A laptop with a Core i7 CPU operating at 2.8 GHz runs the system and its stages, which are implemented using Python 3.8.

Initially, the investigations used the iris image before preprocessing. In this case, we use DenseNet to simulate the eye's iris in its normal size.

Table 2 The specifications of the used datasets

Dataset	CASIA-Iris-Interval	CASIA-Iris-Thousand V4	IIT Delhi	UBIRIS.v1
Number of subjects	249	898	224	241
Samples per subject	20	20	10	10
Number of images	2639	17,938	2240	1876
Image size (pixels)	320 × 280	640 × 480	340 × 280	200 × 150
Image format	JPEG	JPEG	JPEG	BMP

Table 3 The iris image databases' recognition accuracy

Dataset	Before segmentation		After normalization		Use the first level DWT	
	Accuracy	Training time	Accuracy	Training time	Accuracy	Training time
CASIA-Iris-Interval	96%	00.09.08	99%	00.06.08	99%	00.01.07
IIT Delhi	94%	00.06.06	99%	00.05.6	99%	00.01.55
UBIRIS.v1	93%	00.02.36	97%	00.01.31	97%	00.00.55
CASIA-Iris-Thousand	91%	00.41.09	97%	00.24.11	97%	00.12.18

In the second case, the iris image is normalized. The normalizing step sets the radial and angular resolutions to 224. The normalized picture is 224 pixels wide and 224 pixels height, perfect for dense network input.

Case 3: Use the iris picture after DWT and normalization. The picture size changes from DWT 128-by-128 to 64-by-64 to 32-by-32.

There will be a random split of each picture set into training and test data. Images make roughly 80% of training data and 20% of testing data.

This is followed by feature extraction and classification using a pre-trained CNN model. DenseNet was pre-trained to identify 1000 classes [17]. We used the following database types to test the system on all participants in the dataset:

With the same dataset as the class dataset, we pre-trained a DenseNet, using a tuning approach that modifies outputs in the final fully connected layer (FC) based on the class count.

Our study used DenseNet2016 pre-trained with varied epoch counts. The complete training set traversed by the training process is one epoch [18]. The initial learning cycle's learning rate is 0.001. The mini-batch option has a batch size of 32. In computing the gradient of the loss function and adjusting weights, the words mini-batch and subset are interchangeable. Gradient descent iteratively modifies network parameters (weights and biases) [19, 20] (Table 3).

These are the database accuracy and training time before and after segmentation of the ocular image, as well as after utilizing the first level of DWT. Because DWT reduces the picture size while maintaining important features, it reduces training time and increases accuracy. So our method improves accuracy without adding training time.

5 Conclusions

A pre-trained CNN was utilized to build an effective iris recognition system (DenseNet201). Then, an iris image is segmented using a circular Hough transform and normalized using Daugman's rubber-sheet model. The normalized picture is loaded into the pre-trained CNN (DenseNet201). The approach works well on public datasets. We will evaluate the suggested system's performance on other iris datasets and other biometric recognition tasks in the future.

References

1. Thomas T, George A, Devi KI (2016) Effective iris recognition system. Proc Technol 25:464–472
2. Al-Waisy AS, Qahwaji R, Ipson S, Al-Fahdawi S, Nagem TA (2018) A multi-biometric iris recognition system based on a deep learning approach. Pattern Anal Appl 21(3):783–802
3. Liu M, Zhou Z, Shang P, Xu D (2019) Fuzzified image enhancement for deep learning in iris recognition. IEEE Trans Fuzzy Syst 28(1):92–99
4. Sujana S, Reddy VSK (2021) An effective CNN based feature extraction approach for iris recognition system. Turkish J Comput Math Educ (TURCOMAT) 12(6):4595–4604
5. Alaslani MG, Elrefaei LA (2019) Transfer lerning with convolutional neural networks for iris recognition. Int J Artif Intell Appl 10(5):47–64
6. CASIA Iris Image Database Version 4.0 (CASIA-Iris-Thousand). Available from: http://biometrics.idealtest.org/dbDetailForUser.do?id=4
7. CASIA Iris Image Database Version 4.0 (CASIA-Iris-Interval). Available from: http://biometrics.idealtest.org/dbDetailForUser.do?id=4
8. IIT Delhi Database. Available from: http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Iris.htm
9. UBIRIS Image Database Version 1.0. Available from: http://iris.di.ubi.pt/ubiris1.html
10. Sharma VP, Mishra SK, Dubey D (2013) Improved iris recognition system using wavelet transforma and ant colony optimization. In: 5th International conference on computational intelligence and communication networks (CICN). IEEE, New York, pp 243–246
11. Jyoti P, Parvati B, Sandeep KG, Shubh Lakshmi A (2016) New improved feature extraction approach of iris recognition. Int J Comput Syst 3:1–3
12. Ma L, Tan T, Wang Y, Zhang D (2004) Efficient iris recognition by characterizing key local variation. IEEE Trans Image Process 13:739–750
13. Verma P, Dubey M, Basu S, Verma P (2012) Hough transform method for iris recognition-a biometric approach. Int J Eng Innov Technol (IJEIT) p 1
14. Verma P, Dubey M, Verma P, Basu S (2012) "daughman"s algorithm method for iris Recognition-a biometric approach. Int J Emerg Technol Adv Eng p 2
15. Daugman JG (1993) High condence visual recognition of persons by a test of statistical independence. IEEE Trans Pattern Anal Mach Intell 15:1148–1161
16. Srivastava N, Hinton GE, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overftting. J Mach Learn Res 15:1929–1958. ISSN: 1532-4435
17. Huang G, Liu Z, van der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: IEEE Conference on computer vision and pattern recognition (CVPR). Honolulu, pp 2261–2269
18. Al-Waisy AS et al (2017) A multi-biometric iris recognition system based on a deep learning approach. Pattern Anal Appl pp 1–20
19. Syafeeza A et al (2015) Convolutional neural networks with fused layers applied to face recognition. Int J Comput Intell Appl 14(03):1550014
20. Abdulreda A, Obaid A (2022) A landscape view of deepfake techniques and detection methods. Int J Nonlinear Anal Appl 13(1):745–755. https://doi.org/10.22075/ijnaa.2022.5580

IoT-Based Smart Doorbell: A Review on Technological Developments

Abhi K. Thakkar and Vijay Ukani

Abstract Advancement in technology and connectivity has led to the development of the Internet of Things such that in the present era it is now being deployed into daily-use objects. The idea is to control the basic functions and features of these objects through a computing device either automatically or remotely. The class of such objects that come under the umbrella of a home of a person is called smart home devices, and one of them is the smart doorbell. Smart doorbell offers the main feature of security, remote control, and other features depending on the sensors and functionalities it packs. Security from intruders or an unknown person is the primary worry of a homeowner, and hence, smart doorbells can be a solution to ease out this worry. This paper reviews the developments in the field of IoT-based smart doorbell proposed by researchers and also discusses the results, limitations, and possible future developments.

Keywords Internet of Things · Smart doorbell · Smart home devices

1 Introduction

Doorbells have been in use for more than a century. The most basic use of them is to alert the owner of a place about the presence of a visitor or unknown person at their doorstep. Those were primitive times. Over the years, doorbells have evolved from mere electronic devices to sophisticated security systems. From basic door knockers to electronic doorbells to doorbells with cameras and video monitoring to today's smart doorbells with a range of functions, this evolution has gone through several phases.

As a result of advancements in the fields of Internet of Things, artificial intelligence, machine learning, security systems, communication and network technology,

A. K. Thakkar (✉) · V. Ukani
Institute of Technology, Nirma University, Ahmedabad, Gujarat 382421, India
e-mail: 20mced01@nirmauni.ac.in

V. Ukani
e-mail: vijay.ukani@nirmauni.ac.in

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_18

and various other domains, the field of smart security systems and smart home automation has seen multiple advances and research innovations. IoT-based smart doorbell systems are one of the sub-areas that fall under the scope of each of the above-mentioned areas. These systems have evolved over time, adding automation features such as remote monitoring and communication, automatic unlocking, and security features such as visitor warning, one-time password (OTP), biometric security, and so on.

An IoT-based smart doorbell is a concept that combines (but is not limited to) the previously listed domains to create an efficient and intelligent system that is connected to a local network or the Internet and sends notifications to the owner's phone or any other electronic device registered with the doorbell via Web service or application. The smart doorbell is activated in one of the two ways: by pressing a button or by detecting a visitor through an on-device sensor depending on the architecture of the system. It then performs the actions that have been programmed into it.

This review paper would address all of these aspects by providing a brief overview of the various research papers, advances, and recommendations made by researchers and authors with a scientific or technological background, in the hopes of providing potential researchers with enough background to form a scientific opinion and perception.

2 Preliminaries

This section provides a brief description of the key terminologies associated with IoT-based smart doorbell systems as mentioned in the different works of authors described in Sect. 3.

2.1 Internet of Things

The Internet of Things (IoT) is a network of interconnected devices that interact with one another autonomously over the Internet. The IoT ecosystem comprises smart devices that collect and send data and act on that data according to the way they are programmed. By connecting to a gateway or other edge node, the sensors can share the sensor data they collect, which can then be sent to the cloud for analysis or analyzed locally. These devices will sometimes communicate with one another and act on the information they receive.

2.2 Home Automation

Home automation refers to the use of technology to automate or monitor different household functions remotely. It includes connecting all of your home's devices to a single central controlling unit that automates the devices based on the way they are programmed or according to a user's input. These devices are linked through a network, most commonly the Internet, and can be operated remotely. Devices can trigger one another with home automation, and hence, manual operation is not required all of the time. The user can still provide input when it is needed, either directly or from a remote location.

2.3 Video Analytics

Video analytics is a process of monitoring video streams in real time. The program uses video analysis of observed environments to recognize characteristics, activities, or patterns of a particular activity. It can also produce automated warnings and enable historical data analysis to identify trends, patterns, and incidents. For the IoT smart doorbell, it can be a useful tool for recognizing and detecting trespassers, monitoring people or objects, and generating alarms based on certain activities.

2.4 Artificial Intelligence, Machine Learning, and Deep Learning

Artificial intelligence (AI) is a broad term that encompasses all aspects of making machines smarter. The main aim of AI is to teach a computer to think like a human and make decisions on its own. Machine learning (ML) is a sub-field of AI that is often used in conjunction with AI. An AI system that can self-learn based on an algorithm is referred to as ML. ML refers to systems that become smarter over time without human interference. Deep learning (DL) is a form of machine learning (ML) that is used to analyze huge amounts of data. These systems assist a computer model in filtering input data across layers to predict and classify data.

3 Review of Smart Doorbell Systems

This section describes the various aspects of smart doorbells starting from how they have evolved, what kinds of system design and architecture are possible, the connectivity, and hardware technologies. It also states the various techniques employed for the provision of security and then states some current advancements made for the development of IoT smart doorbell.

3.1 Doorbell Evolution

The doorbell is a simple instrument that has been aiding humans for centuries in the simple task of allowing a visitor to inform or alert the owner about his/her presence. As the technology has made advancements, this simple instrument has seen various stages of evolution, and in the present time, it has become complex and advanced with a variety of features to offer. The surge in the field of the Internet of Things (IoT), automation and development of performance efficient devices, and networking technologies has spearheaded this transformation.

Doorbells are now associated with the term “smart” which means that these devices are capable enough to do the tasks which earlier required human efforts. The new age IoT-based smart doorbells are equipped with technology to provide numerous functionalities such as video surveillance, remote access, and face recognition, some of which will be discussed in further sections.

3.2 System Architecture

The smart doorbells are designed to facilitate users with multiple functionalities depending on which system architecture must be chosen. The entire system can be designed using the architecture employed by Thabet et al. [12], Muqeet et al. [6], and Pawat et al. [9], where the central unit is the doorbell device which is responsible for the actions and processing as shown in Fig. 1. The doorbell unit might be configured with other sensors and actuators as required.

Another possible architecture is the one that has a server in the middle to handle requests and data processes. This kind of architecture has been proposed in the work of Jain et al. [4] and Simon et al. [11]. It facilitates the use of advanced algorithms which cannot be deployed directly on the device due to resource constraints and a

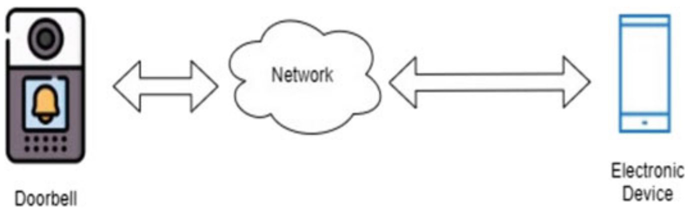


Fig. 1 Generic architecture with doorbell as the central unit

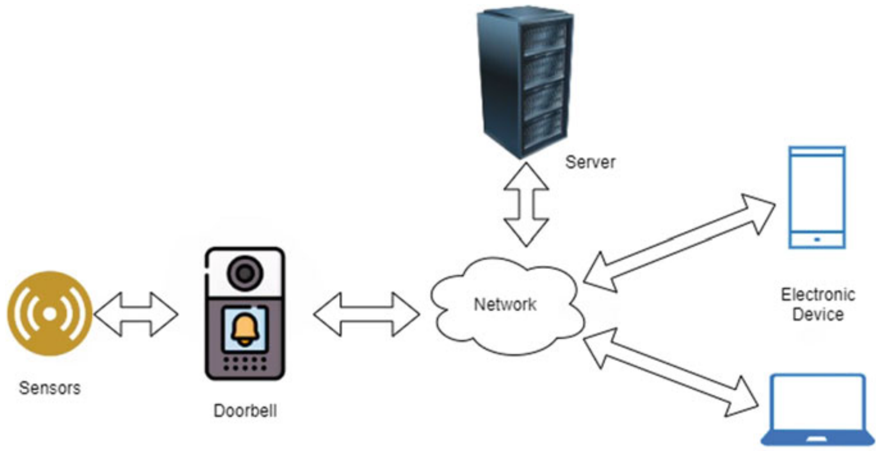


Fig. 2 Generic architecture with physical or a cloud-based server as the central computing unit

central system for managing requests. The server can either be a physical server, data center, or a cloud-based server. Figure 2 shows this generic server-based architecture.

Although the above-mentioned system design and architecture are the two most common methods followed for the development of an IoT-based smart doorbell system, the researchers or developers do not need to limit themselves to these. These architectures might form the basis for the system, and then, a new design can be used such as done by Giorgi et al. [3]. In their work, they have employed multiple AXIOM boards to work in parallel with each other to improve the efficiency of the system.

3.3 Connectivity Technologies

After the system design comes the question about connectivity. The kind of network technology to be employed highly depends on the type of system and is required. For example, if the smart doorbell system is for a small place and the devices connected to the network are mostly stationary, then one of the choices can be to use a wired connection to connect all devices. If the system requires a wireless mode of communication, then the developer can employ the Bluetooth technology or Global System for Mobile Communications (GSM) or WiFi. This also depends on the hardware which whether or not it has support for these networking media.

J. Ding et al., in their paper [2], discuss different types of connectivity technologies for IoT-based smart devices and have provided an excellent classification of such connectivity technologies based on requirements of the system as shown in Table 1. For IoT-based smart doorbell systems, WiFi, Bluetooth/BLE, or other cellular technologies seem to be suitable. Although most modern systems now depend on

Table 1 Summary of different connectivity technologies for IoT systems given by Ding et al. [2]

Category	Use cases	Connectivity technologies
Human-oriented	Smart phone	Legacy cellular technologies, LTE/LTE-A, 5G, WiFi/WiFi HaLow, OWC
Machine-oriented	Monitoring sensors	Bluetooth/BLE, ZigBee, LPWAN, WiFi/WiFi HaLow, OWC
High data rate	Streaming video cameras	LTE/LTE-A, 5G, OWC, WiFi
Long range	UAVs, smart farming sensors	LTE/LTE-A, 5G, LoRa, Sigfox, NB-IoT, LTE-M, WiFi HaLow
Short range	Smart home appliances	Bluetooth/BLE, ZigBee, OWC, WiFi
Low power	Tracking sensors, smart retail sensors	Bluetooth, ZigBee, LTE/LTE-A, 5G, WiFi
Ultra low power	Pollution monitoring sensor	BLE, WiFi HaLow, LPWAN: LoRa, Sigfox, LTE-M, NB-IoT
High mobility	Autonomous vehicles	LTE/LTE-A, 5G
Low mobility	Smart traffic lights	LPWAN, Bluetooth/BLE, ZigBee

WiFi, there are systems that employ Bluetooth technology as well, such as the work proposed by Martinez [5]. Their system employs Bluetooth Mesh Technology which means the nodes in the network are connected in a mesh network using Bluetooth. The smart doorbell uses this technology to send and receive messages inside a building [5]. This is what helped in expanding the reach to the relay nodes. The system performed well at closer distances with fast message transfer but when the distance between the nodes exceeded 10 m, an increased packet loss was observed, which is directly proportional to distance.

The Bluetooth technology has its advantage for small spaces and continuous connectivity but it is also limited by its lack of mobility and necessity to stay in range for effective transmission. WiFi has an advantage over Bluetooth with faster message transfer and long communication range, and if configured with the Internet, then it also enables remote access and control as required by the application.

3.4 Hardware Technology

There are different kinds of connectivity technologies available but it is also important to select which kind of hardware is suitable for the given IoT-based systems from the abundance of technologies available. The ability of the device to perform its task depends on the underlying hardware. Ojo et al. [7] have reviewed different types of IoT devices by categorizing them as low-end devices, middle-end devices, and high-end devices.

Low-end devices have constraints in terms of resources [7] and have RAM capacities equal to or lower than ten kilobytes and processing units of 8-bit and 16-bit architecture, as well as some have a 32-bit architecture. They are used as basic sensing and actuating devices, with low functionalities, and are programmed using low-level languages.

The middle-end devices have an upper hand in terms of resources when compared to low-end devices. They have RAM capacities ranging from ten to hundreds of kilobytes and clock speeds in MHz. They are more capable than low-end devices.

High-end devices are sometimes also called Single Board Computers (SBCs) as they have powerful resources which include a powerful processing unit, a high capacity RAM, high storage volume, and even sometimes a Graphics Processing Unit (GPU [7]). They have support to run lightweight operating systems and are capable enough to perform complex computations such as a machine learning algorithm tentatively and also come with on-the-board connectivity modules such as Ethernet interfaces, Bluetooth/BLE module, and WiFi module.

The most common and frequently used high-end device is a Raspberry Pi called Rpi. This device has been used in plenty of applications one of them being smart doorbell systems, for which variety or literature can be found. The systems proposed by Jain [4], Chaudhari [1], Pawar [9], and Thabet [12] all of them make use of RPi board as the central device for the smart doorbell some of which employ face recognition algorithm which the RPi is capable of running it locally. Smart doorbells are not confined to using only high-end devices. Various sensors and actuators which fall into low-end and middle-end categories are also required for different applications.

3.5 System Features

The purpose of the development of such advance smart doorbell systems is to provide convenience and a sense of security to the owner. Hence, this section states some features that can be incorporated into a smart doorbell by reviewing the work done by other researchers in the same field.

Notification and Alert Notifying the owner about a visitor is a functionality that is a must. It helps the owner to keep track of how many his/her house visited by some person. Simon et al. [11] present a very simple design where when a push-button placed outside the house is pressed thrice the system is programmed to send an SMS to the owner of the house to alert him/her.

A more updated approach is used by Muqeet [6] and Park et al. [8] where along with an SMS being sent there is a camera attached to the system which captures the image and send it to the owner through email. Advancing to this [4, 9, 12] propose a system where the owner is alerted through a notification on an application which may be a Web app or smartphone application running on their device. The owner can also access the images captured and view them through these applications.

Video Surveillance, Two-way communication, and Remote Access If the smart doorbell system is equipped with a camera unit, then apart from taking pictures of the visitor it can also be used to monitor the area where the doorbell is placed. Also if equipped with a mic, a speaker, and in some advanced systems a screen, there can be two-way communication between the visitor and the owner. The system can be linked with an application that provides the functionality of remotely controlling the doorbell devices either to monitor or to allow/deny entry to the visitor.

The system proposed by Thabet et al. [12] provides two of these functionalities. According to their design, the owner can monitor the surrounding area through a video camera by using a Web application. The administrative access through the Web application allows the user to add a previous picture of the visitor into the database to allow future entry. Jain et al. [4] use a more advanced approach where they have an android application connected to the RPi device over the Internet through a server. The owner can monitor through the application, update entries in the real-time database, and also grant access or deny a visitor from entering the home. A future work proposed for this system is to establish two-way communication.

Security Features This is the core feature of an IoT smart doorbell. There are a variety of possibilities for this feature, each having its own advantage and limitation. In [11], the author proposes the most traditional way in which a smart doorbell can provide security. This involves using a keypad to control the lock mechanism through a *password*. This has a major disadvantage of not being autonomous hence failing the entire concept of making *"smart"* doorbells. A concept of *one-time password (OTP)* can be used to automate this process. If the user approves, an OTP is displayed on the screen that needs to be entered by the visitor to unlock the door.

The system proposed by Muqeet [6] takes a newer approach of using a *fingerprint scanner*. This enables automation to a certain extent where all the people that have their fingerprint registered will be allowed to enter the house. The disadvantage of the system is that to add a new visitor his/her fingerprint needs to be added manually. This limitation can be removed if the owner can allow remote registration of the visitor's fingerprint.

Although fingerprints are secure enough, a higher amount of security can be introduced by using *face recognition*. The facial features of every person are unique as well as provide a lot more features compared to fingerprints hence building a more secure system. While [1, 4, 9, 12] propose the use of face recognition in their works, the algorithms used by all of them are different.

In [9], the author proposes the use of the Local Binary Pattern (LBP) algorithm and OpenCV library to develop the face recognition system using Python. The LBP algorithm works by changing the pixel values into 1 or 0 depending on the threshold value to extract the features of the image as shown in Fig. 3. The advantage of LBP is that it is fast and produces efficient results for recognition. OpenCV is an image processing library that is helpful to process the images captured by the camera.

Another approach for face recognition is to use Haar Cascades. This involves the use of Haar-like features to detect a face of a person as shown in Fig. 4. After the detection of the face, the Eigen face algorithm available in the OpenCV library is applied to the image of the face to perform recognition. The author reports a testing accuracy ranging between 85 and 95% and also that their proposed method performed better than other traditional algorithms based on the recognition rate.

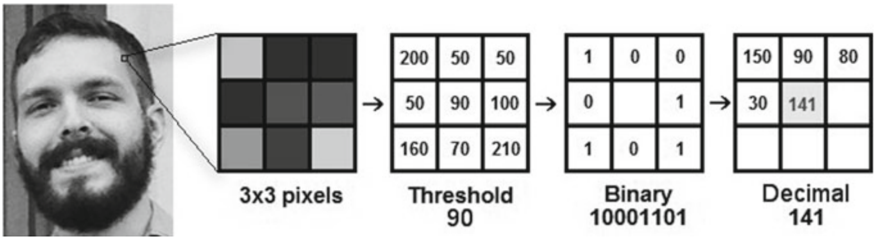


Fig. 3 Local binary pattern

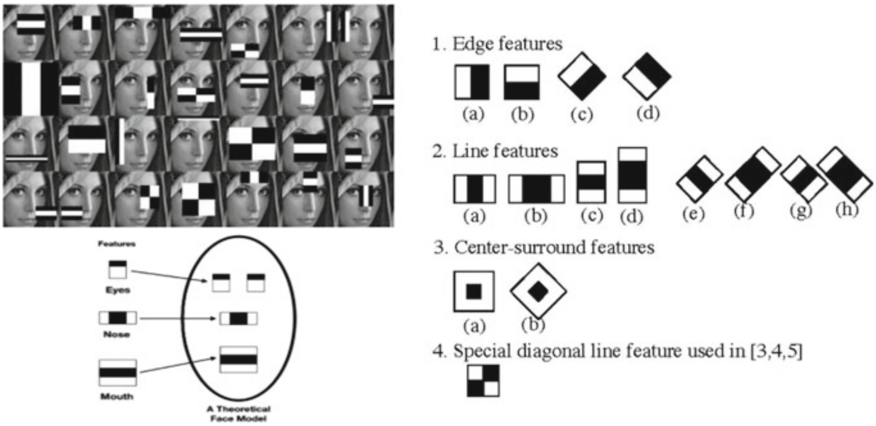


Fig. 4 Haar Cascades

The previous two methods are quite simple and do not require many resources to run. Hence, they can be run on a device such as an RPi itself. There are models based on deep learning which can be employed for more efficient face recognition but the problem arises that generic IoT boards such as RPi do not have sufficient resources to train the model. For this purpose, very high-end devices such as NVIDIA Jetson can be used or the processing of data such as training of model can be done at server side which has enough resources. In [4], a similar design is proposed in which the author has suggested the use of TensorFlow open-source library and Google's FaceNet deep learning model to perform face recognition.

4 Challenges and Future Enhancements

With the study of previous works, it can be deduced that there are many challenges and future enhancements that can be made for IoT-based smart doorbell system some of which are discussed below.

- The systems fail in case of power failure if they are connected to WiFi as WiFi Routers are powered by the main power supply or if they are themselves connected to the main power supply. Hence, solutions such as having backup power through batteries and a network module such as a GSM module through which the device can be connected to the Internet with a power supply from a battery can be thought of.
- There can be challenges related to the connectivity technologies. Samuel [10], in his paper, suggests various issues such as interoperability, maintainability, power consumption, and bandwidth, which need to be taken into account.
- Another addition can be made to establish two-way communication between the owner and visitor with the owner being at a remote place, thus enabling a real-time interaction.
- Adding a feature of suspicious activity recognition and detection of hazardous items.
- Having multiple cameras apart from the one attached to the doorbell module to have a 360° surveillance. This may also enable to have an eye on neighboring areas to help detect suspicious activities.
- Automatically contacting emergency services such as the police in case of intrusion or suspicious activity.

5 Conclusion

In this paper, we introduced the IoT-based smart doorbell systems and discussed the variety of technological aspects involved. The evolution of the doorbell to smart systems, possible systems design and architecture, different connectivity technologies, hardware devices, and features such as surveillance, remote access, and security were discussed by reviewing the works proposed by different researchers. It can be observed that how steadily progress has been made in the field of IoT-based smart doorbell systems with new features added and the old ones updated or swapped out. The paper also discussed some of the challenges that need to be addressed and future enhancements which are possible. The increase in the use of smart devices makes it particularly important that these systems are designed in ways that deliver the services and make the lives of humans better.

References

1. Chaudhari U, Gilbile S, Bhosale G, Chavan N, Wakhare P (2020) Smart doorbell security system using IoT. Technical report, EasyChair
2. Ding J, Nemati M, Ranaweera C, Choi J (2020) IoT connectivity technologies and applications: a survey. IEEE Access 8:67646–67673. https://doi.org/10.1109/access.2020.2985932

3. Giorgi R, Bettin N, Ermini S, Montefoschi F, Rizzo A (2019) An Iris+ voice recognition system for a smart doorbell. In: 2019 8th Mediterranean conference on embedded computing (MECO). IEEE, New York, pp 1–4
4. Jain A, Lalwani S, Jain S, Karandikar V (2019) IoT-based smart doorbell using Raspberry Pi. In: International conference on advanced computing networking and informatics. Springer, Berlin, pp 175–181
5. Martınez C, Eras L, Dominguez F (2018) The smart doorbell: a proof-of-concept implementation of a bluetooth mesh network. In: 2018 IEEE third Ecuador technical chapters meeting (ETCM), pp 1–5. https://doi.org/10.1109/ETCM.2018.8580325
6. Muqeet MA (2019) Fingerprint module based door unlocking system using Raspberry Pi. Sci Technol Dev 8:293–296
7. Ojo MO, Giordano S, Procissi G, Seitanidis IN (2018) A review of low-end, middle-end, and high-end IoT devices. IEEE Access 6:70528–70554. https://doi.org/10.1109/ACCESS.2018.2879615
8. Park WH, Cheong YG (2017) IoT smart bell notification system: design and implementation. In: 2017 19th International conference on advanced communication technology (ICACT). IEEE, New York, pp 298–300
9. Pawar S, Kithani V, Ahuja S, Sahu S (2018) Smart home security using IoT and face recognition. In: 2018 Fourth international conference on computing communication control and automation (ICCUBEA). IEEE, New York, pp 1–6
10. Samuel SSI (2016) A review of connectivity challenges in IoT-smart home. In: 2016 3rd MEC international conference on big data and smart city (ICBDSC), pp 1–4. https://doi.org/10.1109/ICBDSC.2016.7460395
11. Simon OA, Bature UI, Jahun KI, Tahir NM (2020) Electronic doorbell system using keypad and GSM. Int J Inf Commun Technol 2252(8776):8776. https://doi.org/10.11591/ijict.v9i3 (pp 212–220)
12. Thabet AB, Amor NB (2015) Enhanced smart doorbell system based on face recognition. In: 2015 16th International conference on sciences and techniques of automatic control and computer engineering (STA). IEEE, New York, pp 373–377

Development of Student's Enrolment System Using Depth-first Search Algorithm

Ahmed Qassim Hadi, Zainab Adnan Abbas, and Zahraa Mohammed Hilal

Abstract This paper presents a new electronic system that uses technology which is the student enrolment system (SES) used at the University of Technology (Baghdad). The work aims to save time and effort for both the direct admission staff and the student. Application (SES) was evaluated using several variables: (student score, school branch, student desire, role of success in the high school, admission channel, application date and number of seats allocated to each admission channel in each department). The depth-first search algorithm was used to search in data as a tree or graph data structure and speed up the admission process for first-year students at University of Technology (Baghdad). The results obtained from the (depth-first search) algorithm was tabulated using the SPSS statistical programme (statistical version 20) to analyse variance to determine which of the factors most affected the results of student enrolment to the departments. The results showed that the electronic system (SES) included the acceptance of the highest rates in the scientific departments according to the student's desire and within the qualitative capacity of each department. The results showed that the best average was present at the student model 3 (0.218567 ± 0.0044792), which means that the student made the best choices. The results of (ANOVA) showed that when the value of ($P \leq 0.05$), there will be statistical significance, as it was found that the student score ($p = 0.001$), the role of success ($p = 0.003$) and the number of seats available for each department ($p = 0.004$) are among the most influential factors on student acceptance, followed by the student branch ($p = 0.033$) and application date ($p = 0.060$), while the student's desire ($p = 0.549$) and acceptance channel ($p = 0.56$) have the least influence on the student's acceptance.

A. Q. Hadi (✉) · Z. A. Abbas · Z. M. Hilal
University of Technology, Baghdad, Iraq
e-mail: ahmed.q.hadi@uotechnology.edu.iq

Z. A. Abbas
e-mail: zainab.a.abbas@uotechnology.edu.iq

Z. M. Hilal
e-mail: zahraa.m.hilal@uotechnology.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_19

Keywords Depth-first search algorithm · Student enrolment system · Database schema · Graph data structure

1 Introduction

The internet was developed rapidly, and information system over web is quickly increased at unmatched speed [1]. Governments, institution and companies spread their information and manage it over web, which is an important direction. In the websites construction, manage, distribute, mine and utilize information in efficient way are important [2]. While the internet markets are flourish, the concept of web-based applications became more important, processing business and became a method to increase the markets and improve incomes for institutions. Building systems over web may help governments, institution and companies fast set up complicated systems in cheaper operation [3]. Student enrolment system is the first step for students who want to complete their university studies in the Iraqi's universities and institutes to obtain a bachelor's degree or technical diploma. Since 2017, the University of Technology has used its own system to accept new students in nine of its sixteen departments, and this system is the student enrolment system, that uses depth-first search algorithm. There are many researchers who have developed applications that use the DFS algorithm.

Foteini G. and Ioannis H. developed an automatic marking system which assists to lessen the time required in marking and utilize this time effectively for more innovative work. The framework upholds programmed marking of student answers to interactive exercises concerning blind search in DFS. The outcomes demonstrate a decent arrangement between expert and system marking. Additionally, towards the finish of the test, students were approached to finish an internet-based poll about the framework, and it shows that a large portion of the students gave positive reactions [4].

Naoufel W. and Faouzi K. created decision support system (DSS) for student advising. The framework intends to give students with a mechanized programme scheduling and arranging service that best accommodates their profiles while meeting scholarly prerequisites. They found that the DSS was a successful embedding of the essential guidelines, with a systematic and exhaustive search of the academic plans throughout a period scale that ranges from the next term up to graduation [5].

2 Practical Part

2.1 Enrolment System

The direct admission staff in the Registration and Student Affairs Department had to create an electronic system that allows students to apply for admission to the nine departments according to instructions and controls set by the university, in terms of the total Student Competitive "it's calculated through a mathematical equation developed by the university" which it is the student score, school branch "Applied, Biological and Scientific branch", the student's desire, the role of success in high school, acceptance channel and the number of seats assigned to each admission channel in each department in the university. This system will reduce the efforts and load of the staff of the direct admission staff in the Registration and Student Affairs Department by providing simple interfaces for all possible users.

After the results of the high school announced, the university opens the student enrolment system to receive electronic forms for students applying to the university, students start applying by entering the system, and the application involved several steps:

- Fill in the required fields to enter the system "exam number, password and the activation code";
- The entered data will be matched with the system's database, in case that the entered data is correct, the electronic form will appear with the student's basics information from the database on it "student's name, gender, directorate of education, exam number, high school branch and the name of the school";
- Student writes his options in sequence from (1–9), he is not required to write all options, but it is required to write at least one option or leave the system and not submit;
- After completing the form, the student can print the form or save it and leave the system.

The student score is calculated through a mathematical equation developed by the university; the system calculates total competitive by the following operations:

1. Islamic education degree "for Muslim's students" $= A$;
2. Arabic degree $= B$;
3. English degree $* 4 = C$;
4. Mathematics degree $* 8 = D$;
5. Physics degree $* 8 = E$;
6. Biology degree "for biological and scientific students" $* 4 = F$;
7. Economy degree "for applied students" $* 4 = F$;
8. Chemistry degree $* 4 = G$.

Finally, the student score can be calculated by this equation:

$$\text{total competitive} = \frac{(A + B + C + D + E + F + G)}{30}$$

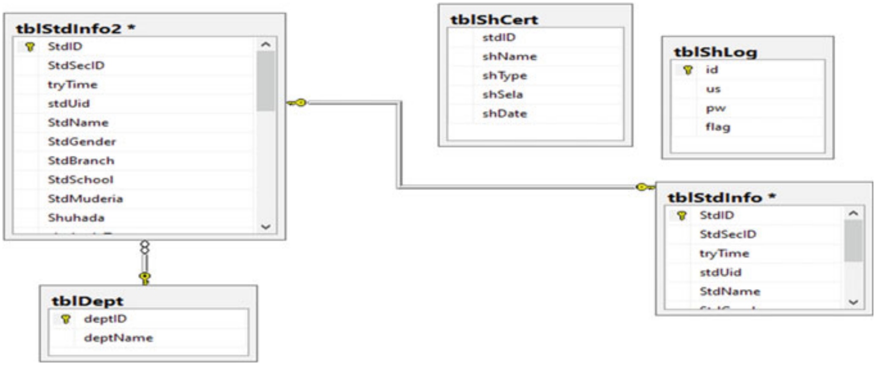


Fig. 1 Database schema

2.2 Database Design

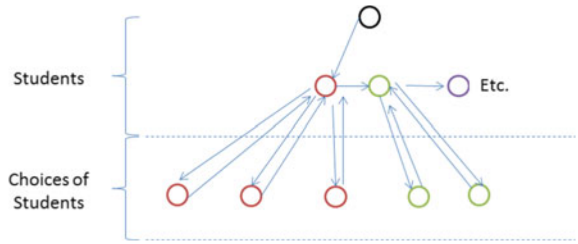
A database is an information structure that stores coordinated data [6]. Most databases contain many tables, which may each contain many fields [7]. Simply, it is collection of coordinated data, normally as a set of related lists of entries. The information is regularly coordinated, so it is effectively accessible. Coming up next are instances of database that we use frequently (address book, dictionary and telephone book) [8]. In this system, we used SQL server since it is actual quick, dependable, easy to use, free to download. Figure 1 shows the database schema.

This database schema of system consists of five main tables. Table "tblStdInfo" contains information about all students. It is connected to the table "tblDept" which they belong to. Each record in table "tblStdInfo" connects to tables: "tblDept", "tblStdInfo", "tblShCert" and "tblShLog"; all of these tables are related by one to many relationship. Table "tblStdInfo" contains information about all students. It consists of 18 fields. These fields are store student's ID, students name, admission channel, students score, student's choices and a date time label. The primary key for this table is "StdID" which is automatically incremented.

2.3 Depth-first Search Algorithm

Depth-first search algorithm is used for searching data in graph or tree structures [9]. It is selecting some arbitrary node for start searching as a root node in the case of a graph, and before backtracking, it explores each branch as far as possible along the idea is to begin searching from the root node, sign it and move to the next unmarked node; this loop continues until all nodes signed, then backtrack to search for other unsigned nodes. Finally, the nodes are printed in the path [10]. The idea for using depth-first search algorithm in this system is to find the first choice from student's choices while ensuring acceptance of the highest rates in the scientific departments

Fig. 2 Depth-first search algorithm



according to the student's desire and within the specific capacity of each department; Fig. 2 shows how the algorithm works in system.

2.4 Statistical Analyses

The student's data was analysed statistically, by using analysis of variance (ANOVA). The specific values of results were done with the SPSS (statistics version 20). The results of the analysed data were estimated significant when the $P \leq 0.05$ [11].

3 Results

3.1 Data Test

In this process, the data tested by executing all processes to all of data. This process was depicted in Fig. 3. The first step was read descending by student score, the second step was looping in the student's choices, if student's score greater than or equal to the lowest limit of the department, then it goes further to the third step, else return to the next choice, the third step was read numbers of an available seats for that department, if there is a seat available, then it goes to the fourth step, else goes to the next choice, the fourth step was accepting the student in department, and that was the last step for the data test [12–15].

Design of interface

Design centres around expecting what users may have to do and guaranteeing that the interface has components that are not difficult to get to, straightforward, reasonable and use to work with those activities [13].

The implementation of SES interfaces. Figure 4 shows the main page "Home page" of the system. It contains textboxes "Enter exam number", "Enter password", "Enter the activation code (Captcha)" and "sign in" button.

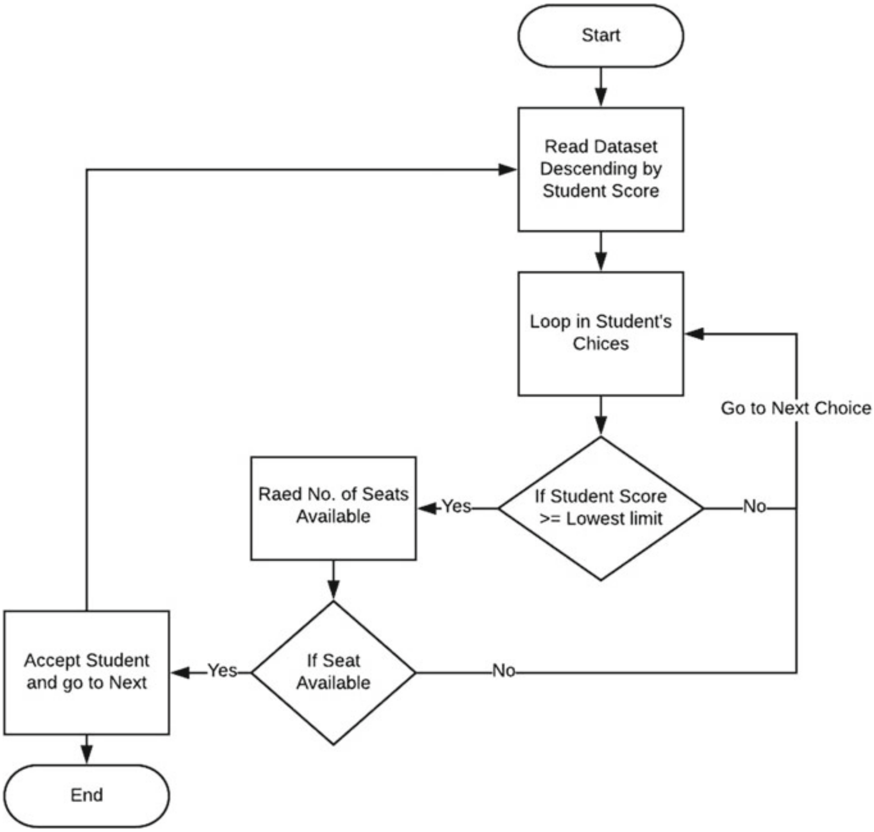


Fig. 3 Students enrolment system using depth-first-search algorithm

Figure 5 shows the page for "student page", it appears when the student entered correct information, this page shows student's information, and the student must choose his admission channel and fill-in his choices and submit.

Figure 6 shows final page as a report; the student must print this page and save it.

3.2 Analyses Static

Table 1 describes results of the mean SES; the relation between mean of ratio and model of students is displayed in Fig. 7.

The results of Table 1 show that the best mean was present at the student model 3 (0.218567 ± 0.0044792), which means that the student made the best choices. This can be explained by the student arranging the area (0.001, 0.003 and 0.004), respectively, has less factors in a better way compared to the rest of the students.

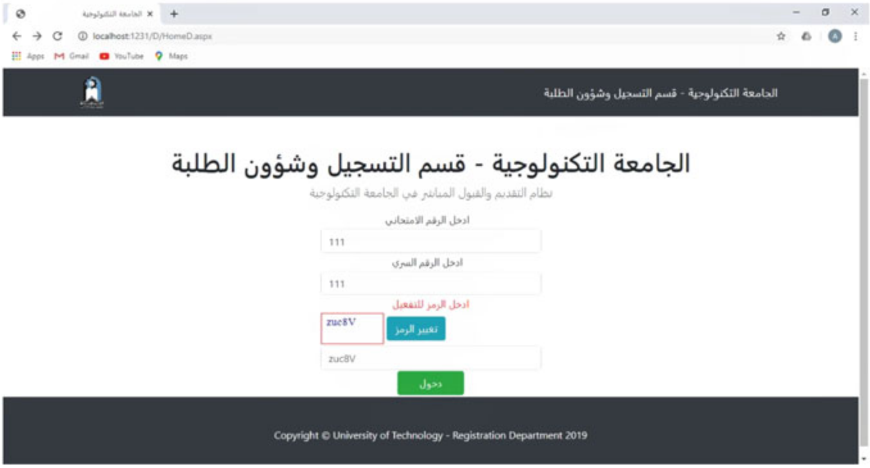


Fig. 4 Home page

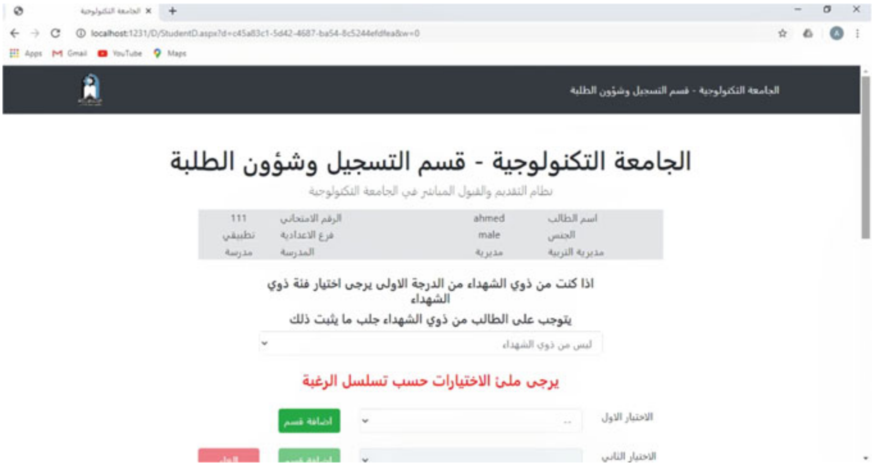
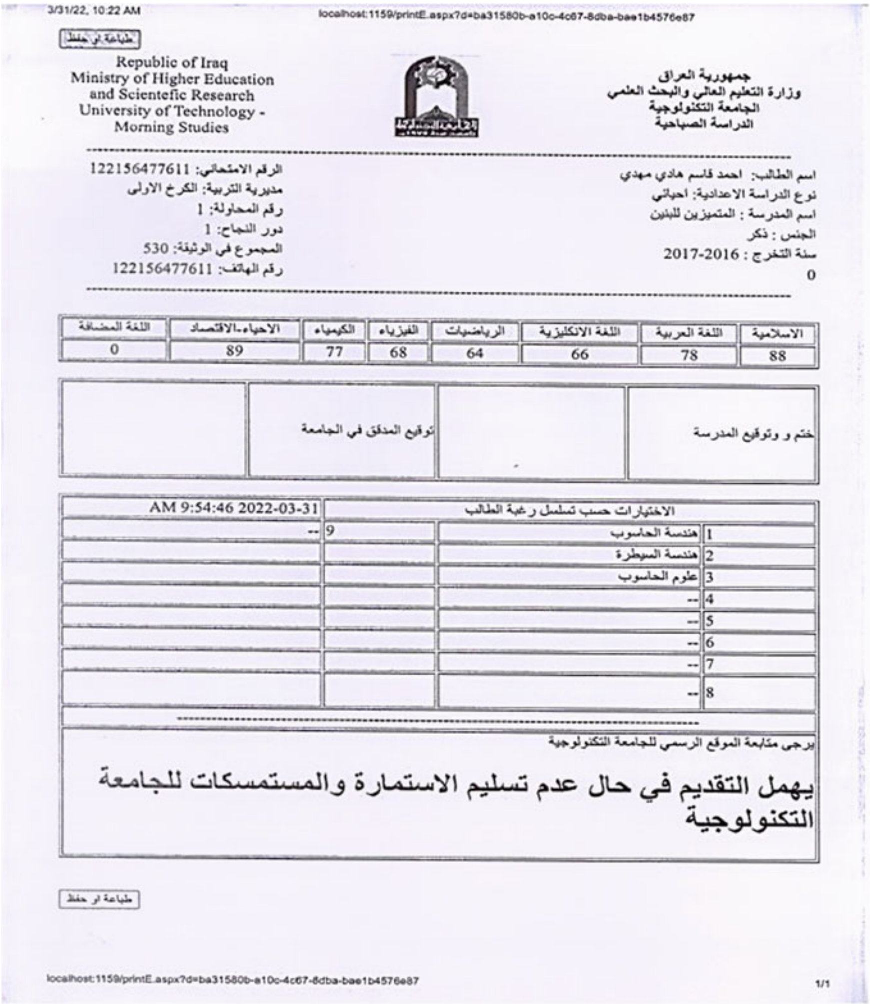


Fig. 5 Student's page

Table 2 offers the ANOVA static analyse for all model of students under all variables. There is variation in the mean value between variable of models of students greater than expected. The *P*-value of student score, role of success in the school and number of seats allocated than 0.05.



3/31/22, 10:22 AM localhost:1159/printE.aspx?d=ba31580b-a10c-4c67-8dba-bae1b4576e87

طباعة او حفظ

Republic of Iraq
Ministry of Higher Education
and Scientefic Research
University of Technology -
Morning Studies

جمهورية العراق
وزارة التعليم العالي والبحث العلمي
الجامعة التكنولوجية
الدراسة الصباحية

اسم الطالب: احمد قاسم هادي مهدي
نوع الدراسة الاعدادية: احيائي
اسم المدرسة : المتميزين للبنين
الجنس : ذكر
سنة التخرج : 2017-2016
0

الرقم الامتحاني: 122156477611
مديرية التربية: الكرخ الاولى
رقم المحاولة: 1
دور النجاح: 1
المجموع في الوثيقة: 530
رقم الهاتف: 122156477611

الاسلامية	اللغة العربية	اللغة الانكليزية	الرياضيات	الفيزياء	الكيمياء	الاحياء-الاقتصاد	اللغة المضافة
88	78	66	64	68	77	89	0

ختم و وتوقيع المدرسة		توقيع المدقق في الجامعة	

	الاختيارات حسب تسلسل رغبة الطالب		2022-03-31 9:54:46 AM
1	هندسة الحاسوب	9	--
2	هندسة السيطرة		
3	علوم الحاسوب		
4	--		
5	--		
6	--		
7	--		
8	--		

يرجى متابعة الموقع الرسمي للجامعة التكنولوجية

يهمل التقديم في حال عدم تسليم الاستمارة والمستمسكات للجامعة التكنولوجية

طباعة او حفظ

localhost:1159/printE.aspx?d=ba31580b-a10c-4c67-8dba-bae1b4576e87 1/1

Fig. 6 Printed page

4 Conclusion

The goal of the system is to develop of student's enrolments system using depth-first search algorithm to save time and effort for both the direct admission staff and the student. The depth-first search (DFS) algorithm was used to search and speed up the admission process for the enrolment at University of Technology, and the results obtained from the (depth-first search) algorithm were tabulated using the SPSS statistical programme to analyse variance to determine which of the factors most affected the results of student enrolment to the departments. It was found

Table 1 Mean and description of the DAS for the all samples

Samples	Std. deviation	95% confidence interval for mean		Maximum
		Lower bound	Upper bound	
1	0.0032512	0.025724	0.041876	0.0362
2	0.0010817	0.209613	0.214987	0.2135
3	0.0044792	0.207440	0.229694	0.2220
4	0.0013229	0.099214	0.105786	0.1040
5	0.0110151	0.051970	0.106696	0.0900
6	0.0070946	0.122043	0.157291	0.1460
Total	0.0695605	0.096436	0.165619	0.2220

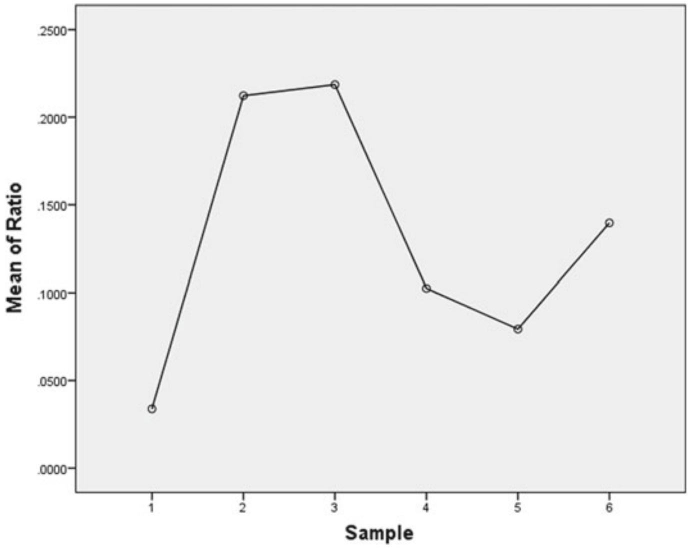


Fig. 7 Relation between mean of ratio and sample

Table 2 Statically results analyse ANOVA of all variables

Source	Adj SS	Adj MS	F-value	P-value
Student score	0.897523	0.179505	82.35	0.001
School branch	0.005527	0.002763	1.27	0.033
Student desire	0.000229	0.000115	0.05	0.549
Role of success in the school	0.838216	0.167643	75.67	0.003
Admission channel	0.004926	0.002463	0.08	0.56
Application date	0.000374	0.000187	1.11	0.04
Number of seats allocated	0.000470	0.000230	60.30	0.004

in ANOVA that the student score ($p = 0.001$), the role of success ($p = 0.003$) and the number of seats available for each department ($p = 0.004$) are among the most influential factors on student acceptance, followed by the student branch ($p = 0.033$) and application date ($p = 0.060$), while the student's desire ($p = 0.549$) and acceptance channel ($p = 0.56$) have the least influence on the student's acceptance.

References

1. Zhao L, Du S-F (2010) Design and implementation of j2EE-based web website content management system. In: 2010 International conference on E-Product E-Service and E-entertainment (ICEEE). IEEE, New York
2. Fang X, Min D (2007) Development and application of content management system. J Xiaogan Univ 3:020
3. Liu B et al (2010) Website building systems design and implementation based on content management. In 2010 International conference on computer, mechatronics, control and electronic engineering (CMCE). IEEE, New York
4. Grivokostopoulou F, Hatzilygeroudis I (2013) An automatic marking system for interactive exercises on blind search algorithms. In: International conference on artificial intelligence in education. Springer, Berlin, Heidelberg, pp 783–786
5. Werghi N, Kamoun FK (2010) A decision-tree-based system for student academic advising and planning in information systems programmes. Int J Bus Inf Syst 5(1):1–18.
6. Deitel P, Deitel H (2007) Internet & world wide web: how to program. Prentice Hall Press
7. Rainer RK et al (2013) Introduction to information systems. Wiley
8. Official site of School of Computer Science, Carleton University. http://moveableonline.com/blog/2013/10/29/7-advantagesusing-cms-run-site/. Date of access: 06 June 2020
9. Russel S, Norvig P (2001) Artificial intelligence: a modern approach. Prentice Hall
10. Limon Y, Barcenas E, Ben'itez-Guerrero E, Medina MA (2018) Depth-first reasoning on trees. Computación y Sistemas 22(1):189–201
11. Mohammed RA, Al-Zubidi AB, Attallah MS (2019) Analyzing mechanical properties of silica gel filled with carbon fiber into polymeric composite. J Univ Babylon Eng Sci 27(3):224–237.
12. Arwan A, Sagita D (2018) Determining basis test paths using genetic algorithm and J48. Int J Electr Comput Eng (IJECE) 8(5):3333–3340
13. Garrett J (2011) Elements of user experience: user-centered design for the web and beyond, 2nd edn. New Riders, USA, p 191p
14. Obaid AJ, Chatterjee T, Bhattacharya A (2020) Semantic web and web page clustering algorithms: a landscape view. EAI Endorsed Trans Energy Web 8(33)
15. Gadicha AB et al (2021) J Phys: Conf Ser 1963:012141

Merging of Internet of Things and Cloud Computing (SmartCIOT): Open Issues and Challenges

Isha Sharma, Prabhsharan Kaur, Pankaj Kumar, and Sheenam

Abstract In the field of computer science, nowadays, there is a major drift in the field of the Internet of things and is rightly labeled as a smart revolution combining the Internet and things around the world. Using IoT, it is possible to allow connections between billions and trillions of devices and things with the help of the Internet and thereby allowing for the exchange of information and making the daily life of human beings easier and smarter and thereby improving the quality of service. Cloud computing on the other side is handy and scalable access to the network that allows access to on-demand services and allows computing resources to be shared those accounts for dynamic information integration from different sources of data. Merging cloud computing with IoT calls for open questions relating to the merger and confrontations that seek attention and discussed to allow for the successful merger of the Internet of things and cloud. Both the technologies are interdependent for their further growth and publicity. Cloud computing offers a huge number of computing resources that can be used by IoT, on the other hand, cloud computing allowing IOT to use the resources available on cloud accounts for dynamic and distributed data integration. The objective of this paper is to give an informative overview of the merits and demerits of the merger of smart devices and cloud computing. The paper also addresses the challenges faced by allowing for the merging of cloud computing and IoT. In the end, the future research ideas about open issues of the merger resulting in a new paradigm SmartCIOT (cloud-based IoT) are also discussed.

Keywords Internet of things · Cloud computing · Cloud services · IaaS · PaaS · SaaS

I. Sharma (✉) · P. Kaur · P. Kumar · Sheenam
Chandigarh University, Mohali, India
e-mail: ishasharma211@gmail.com; isha.cse@cumail.in

Sheenam
e-mail: sheenam.e6717@cumail.in

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_20

1 Introduction

The paradigms IoT and cloud computing are not new areas that need any introduction. The paper is focused on merging these existing paradigms, thereby increasing the power of these technologies. The merging of IoT and cloud computing calls for improvements too. Before the merger, i.e., SmartCIOT, one must understand the technologies independently also.

Cloud computing is the combination of different various other services, e.g., software services, grid services, utility services, virtualization, and networking services [1, 2]. Cloud computing is, therefore, not a new technology but was the one that evolves and period [1]. The roots of cloud computing were found in the area of software services, networking services, grid computing, virtualization, and distributed systems [2]. Cloud computing thus gives parallel computing infrastructure that allows sharing of different computing services and resources with the help of the Internet [3, 4].

IoT, on the other hand, is a dynamic interface that allows the objects around the globe to be connected to a networking domain by adding the intelligence domain to the objects connected via public network, i.e., Internet [5]. The popularity of IoT is because allows the interconnection of things around the world with the help of the Internet without or less human intervention [6]. The objects in IoT have less storage capacity and computing requirements [7]. It would not be wrong to tagging both IoT and cloud computing as the new smart generation technologies IOT technologies are more focused on supporting diversity and allowing different things to be connected in networking scenarios, whereas cloud computing services are more focused on supporting interoperability.

The purpose of this chapter is to give an informative overview of the merger of cloud computing services and IOT services. An attempt was made with this research to examine the merits and issues relating to this merger. An effort has been made to identify and quantify the challenges faced during the implementation. The research paper is outlined into different sections as follows: Sect. 2 of the research paper is discussing the basic introductory concepts about IoT and cloud computing and gives the road for cloud combine with IoT infrastructure. Section 3 discussed the merits of merging IoT and cloud. Section 4 presented the IoT and cloud integration architecture. Sections 5 and 6 discussed the open issues and challenges faced by this smart next-generation merger of these two top technologies. The research paper is finally concluded in Sect. 7.

2 Concepts About IoT, Cloud Computing, and Cloud-Based IoT

The following sections discussed the basic introductory concepts about IoT and cloud computing and give the road for cloud combine with IoT infrastructure.

2.1 Internet of Things

It is defined as a smart approach in this modern era that allows all the physical things to be converted into smart things that support smart services. IoT architecture is made up of different components that are connected in such a manner that they reforming a network in which the physical objects are communicating and actively participating ensuring smart services to the user of the objects [8]. The main components of IoT are sensors and smart devices. The physical objects are not limited to electronic devices that are used in daily lives and also include things like clothes, machine parts, assemblies, landmarks, and monuments even food items. All these IoT devices are tracked, monitored, and counted to account for less loss and cost [9].

The origin of the IoT concept was traced back to 1990 and was first coined by Ashton Kelvin [10]. He has defined IoT as the technology that can change the way the world is using the Internet and hereby will be adding more power to the Internet by using the Internet in converting dumb physical devices into smart devices. In 2005, the International Telecommunication Union (ITU) [1, 11, 12] gave the formal presentation of the Internet of things. As per ITU (2012) [13], the Internet of things is such a technology that allows both real and no real things around the world to be connected, thereby helping in the dissemination of data. This sharing of data is due to the power of a public network named the Internet.

Although IoT supports various smart applications using the Internet, the field is also subjected to face many issues that sometimes behave as the obstacle in the field of the implementation of IoT. IoT is facing the challenges of data storage, scalability, heterogeneity of the things, and energy efficiency concerns [8].

2.2 Cloud Computing

National Institute Of Standards and Technology (NIST) has defined cloud computing as a ubiquitous model that enables on-demand access to the network and allows sharing of a large number of resources like storage applications, servers, networks, and other services. All these resources are computing configurable resources.

Cloud computing consists of four types of deployment models and three different service models mentioned as follows:

- Public versus private cloud
- Hybrid cloud versus community cloud

In public cloud, the resources are shared among the users with the help of Internet and this type of cloud are owned by profitable organization, e.g., Amazon EC2 [9].

Private cloud belongs to a single organization to provide services of a particular type to the users. These types of cloud are coupled with tight constraints in the domain of security and control, e.g., private cloud supported by Microsoft.

The combination of public and private clouds is a hybrid deployment version [10]. The new version is there to allow advantages of both clouds and hence, ignore the limitations of private and public models.

Community cloud is dedicated to a group of users having the same needs, i.e., belonging to users of the same community or organization.

The three different types of services models offered by cloud computing are:

- Sharing of the software services among users of the Internet falling in the category of software as a service model (SaaS), e.g., Google applications.
- Sharing of infrastructure including hardware devices, servers, and storage are framed as infrastructure as a service model (IaaS), e.g., cloud services given by Amazon Web Services
- Sharing of deployment and integrated platform for building, deploying, and testing software are falling in the category of platform as a Service model, e.g., Microsoft Azure [1, 14]

3 Smart Combination of Cloud Computing and Internet of Things (SmartCIOT)

Cloud computing and the Internet of things are next-generation technologies that are now in the developing stage. Both the technologies have their own set of distinct features.

Cloud Computing is a huge network with infinite storage and computation resources. The environment supported by the cloud is robust and flexible which allows data available from several data sources to be integrated dynamically [15].

On the other side, IoT is a smart approach in which a huge number of physical objects are communicating with one another using the power of a global network, i.e., Internet. The devices in IoT are of distributed nature and often suffer from the problems of less storage and limited processing capacities. The devices also suffer from issues relating to security, privacy, reliability, and performance [15, 16].

Both the technologies cloud computing and IoT are being merged to change the environment of how networking services are offered to the distributed users [7].

The differences between the two technologies serve the basis of the speculation made by the researchers but are merged to provide a cloud-based IoT platform, thereby permitting smart use of information, infrastructure, and applications in a cost-effective manner. The difference between the two developing technologies is the main reason researchers have proposed their merger [15]. The difference between cloud computing and IoT is depicted in Table 1.

Table 1 Difference between IOT and cloud computing

Elements	Cloud computing	Internet of things
Features	Ubiquitous technology	Pervasive technology
Resources	Virtual objects	Real-world objects
Storage capabilities	Limitless storage capabilities	Limited/fewer storage capacities
Connection	Service sharing using the Internet	Uses Internet for connection purpose
Big data	Manage big data	Source of big data

4 Merits of SmartCIOT

Because the Internet of things is facing issues related to limited storage, limited processing power, and suffers from issues related to reliability, security, and performance; the merger of cloud and IoT is the best way to solve the maximum issues [17]. Cloud, too, is going to be benefitted from IoT since the merger will expand the boundary of cloud to real-world things making cloud computing a dynamic technology and giving new services to billions of things in this real world [7, 15]. Using cloud computing, it is possible to provide services at a low cost to the end-users [18].

The merits of SmartCIOT are as follows:

- **Improved processing capabilities**

  As we know that IoT devices have limited efficacy of processing capabilities that do not allow the processing of complex data in real-world scenarios. The merger of IoT with the cloud will help in providing unlimited virtual processing capacity of cloud to IoT, thereby improving the processing capacity of the merged model [19].

  Further addition of the data-driven decisions and predictive algorithms in IoT will also help in getting better cost and reduces risk [7].

- **Sharing of application and data**

  Communication is way, too, improved in cloud-based IoT platforms. The two important features of cloud-based IoT are application and data sharing. IoT devices are everywhere and the cloud provides the interface to manage and track everything connected via the Internet [18]. The merger of IoT and cloud will help in better communication employing real-time data tracking and also helps in monitoring real-world objects dynamically [18, 20].

- **New improved features**

  Cloud computing is merged with IoT to solve the issues of IoT related to the areas of reliability, security, and privacy. Due to the merger, it is now possible to use and access real-world objects with less deployment price [11, 21].

- **Novel models**

The cloud-based IoT merger has resulted in new models. The same are listed below [1, 6]:

- Ethernet as a service (EaaS) [15]
- Sensing as a service (SaaS)[1]
- (Identity and policy management as a service (IPMaas) [15]
- Sensing and actuation as a service (SAaaS) [1]
- Database as a service (DBaaS) [15]
- Data as a service (DaaS) [13, 15]
- Sensor as a service (SenaaS) [15, 22]
- Sensor event as a service (SEaaS) [1]
- Internet of things as a service (IOTaaS) [23].

5 Structural Diagram of Smart Merger, i.e., SmartCIOT

The structural diagram supporting smart merger discussed in the above sections is given in Fig. 1 where the network layer of IoT architecture [24] is replaced by the cloud computing layer.

The sensing layer helps in identifying real-world things and collect information from the surrounding environment. The purpose of the intermediate layer is to act as a transporter of the information captured by sensors to the server either on a cloud platform or an Internet platform. The various application-oriented services are supported and provided with the means of the application layer [6, 26].

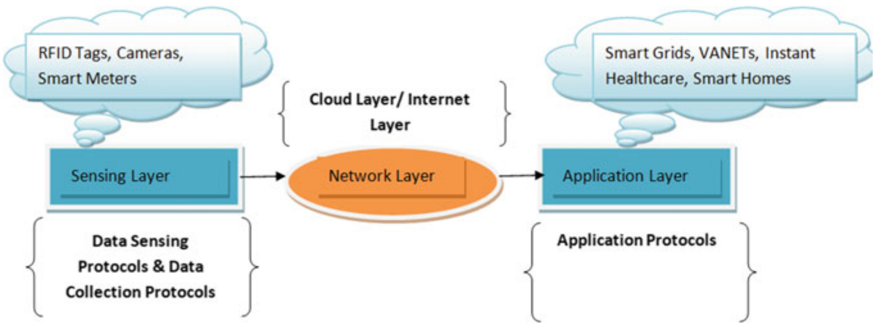


Fig. 1 Structural diagram of smart merger SmartCIOT [25]

6 Applications of Smart Merger SmartCIOT

The SmartCIOT has resulted in new applications and services. The following section gives a brief about applications that are supported by cloud supported IoT paradigm [2].

- Environmental monitoring: Applications that fall under this category are monitoring of water quality, pollution sources, and air quality checks. The combination of cloud and IoT has resulted in a high-speed network information system that can be used to give real-time statistics about the environment.
- Instant healthcare support: Applications that fall under this category are hospital management via smart sensors and drug control. The target of this merger in the healthcare industry is to improve and innovate health care [27].
- Smart innovative homes: Applications that fall under this category are smart home surveillance, home automation, baby monitoring, smart meter reading, and energy saving. The merger of IoT and cloud has opened a path for automation of the home.
- The smarter flow of goods in the logistic industry: Application that falls under this category due to merger of cloud and IoT is smart logistic control that allows and helps in tracking of goods in real time between consumer and producer.
- Smart GPS: Applications that fall under this category are real-time traffic notification and prediction, sensing of the remote vehicle in case of theft. All this is possible due to the merger of cloud computing into a GPS add-on with vehicles that will behave as real things due to IoT.
- Security surveillance: Applications that fall under this category are wireless video capture TV and theft detection systems. The cloud interface provides the platform to store, manage, and process the motion data captured by smart cameras to provide and ensure security.

7 Challenges Faced in SmartCIOT

Although the smart merger discussed in the above sections is quite advantageous and comes up with numerous advantages but in reality, the merger is also subjected to lots of confrontations and issues that seek attention in the context of SmartCIOT and should be brought out in open.

The following are the challenges the innovative merger is facing that needs to be addressed:

- Handling of big data

  IoT devices generate billions of big data every day and because of this huge amount of real-time data that was estimated to be more than 60 billion data units by 2020, there was a need to manage this huge data. Management of this big data involves accessing it, storing it, transporting it, and processing it too. It would be an apt statement that IoT is generating a huge amount of real-time data, and cloud computing allows management of this generated big data for a future period [2]. Management and handling of this data are the main issue of the cloud-based IoT paradigm, and there is a need to find an optimal management solution for big data. Apart from big data management, maintaining data validity is also important to retain the quality of the data.

- Security and privacy.

  Using this platform of integrated cloud and IoT, it is possible now to store the data collected from the real-world entities involved in IoT to the server. Due to the availability of this huge data on the server, the data is also subjected to the problems of privacy and security. Issues like authorization of user allowed accessing big data, trusting in a service provider; service level agreements are still sensitive domains that need to be addressed to make the integration of cloud and IoT the successful one [15, 19]. Leakage of sensitive information may also occur. Cryptology algorithms are even not capable of ensuring privacy in this case [19]. Due to the distributed nature of the integrated paradigm, the merged interface is also prone to several SQL injection attacks and vulnerabilities [19, 28].

- Performance

  High bandwidth is required in the case of the cloud to hold a vast total of information generated by IoT. This directly affects the performance of the network in processing the data being transferred [19]. The slowdown of the cloud in processing the data and transferring the data to IoT devices will affect the timeliness feature of IoT, and hence, keeping the performance upgraded is still an important issue [28].

- Heterogeneity

  The heterogeneous nature of IoT devices is also the main challenge that cloud-based IoT faces. The devices used in IoT are real-world objects and are of different types that are using different operating systems and are used for different applications. This heterogeneity is the main issue the cloud computing technology suffers from [19]. The solution to this is to use multiple cloud interfaces, but this solution is also subjected to be dependent on multiple providers for the same [29, 30].

- Large scale

  The smart merger SmartCIOT allowed the development of new applications that will help in merging and processing the data extracted from the real world into real-world objects used in IoT. This further incurs using multiple devices in billions that are located and distributed across the world. This distribution on such a large scale causes various issues and difficulties [19].

- Legal issues

  The integrated cloud and IoT are also facing issues related to the legal aspect.

8 SmartCIOT Open Issues and Future Guidelines

This section will span open issues and futuristic guidelines related to SmartCIOT.

- Standardization

  The SmartCIOT interface lack standards to be enforced [19]. Although different standards are proposed for SmartCIOT, still more standardization protocols, application programming interfaces are required to allow different IoT devices with cloud service [2, 20].
- Protection mechanism

  Protecting the real-time data generated on a timely basis is the most sensitive area and issues relating to the SmartCIOT [2, 30]. Providing authorization rights, dealing with hackers, access to sensitive data is the new research direction in this direction.
- Fog Computing

  Fog computing is the latest trend and buzzword nowadays and is rightly said to be improving the cloud computing domain by extending it. Fog like cloud computing provides services to the end-users [7]. Fog computing is the latest research buzzword nowadays that is giving extension and edge to the cloud infrastructure.
- Big data

  This SmartCIOT has resulted in a huge amount of big data that is originating daily from different locations and different sources in real time. Hence, keeping the data stored with all the real information intact is a big open issue and calls for further research [19, 31].

9 Conclusion

The IoT is becoming the latest buzzword these days. It is such a computing service that is generating an enormous amount of data. Internet of things is having less capacity in the areas of power, storage, and processing. IoT faces certain issues in the field of reliability, performance, privacy, and security. The merger of IoT with the cloud is, therefore, advantageous in terms of dealing with challenges faced. The requirement behind SmartCIOT was discussed in the above sections. The paper also presented the architecture developed due to the merger of the latest technologies with a focus on challenges the merger is facing and open research guidelines. In the future, more improved architecture is required to overcome the challenges the merger is facing.

References

1. Gai K (2012) Towards cloud computing: a literature review on cloud computing and its development trends. In: 2012 Fourth international conference on multimedia information networking and security, pp 142–146
2. Botta A et al (2014) On the integration of cloud computing and internet of things. In: 2014 International conference on future internet of things and cloud, Barcelona, pp 23–30
3. Buyya R et al (2009) Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility. Futur Gener Comput Syst, pp 599–616
4. Armbrust A et al (2010) A view of cloud computing. Commun ACM 53(4):50–58
5. Altakrouri B et al (2010) Supporting interaction with the internet of things across objects, time and space
6. Lee I, Lee K (2015) The Internet of Things (IoT): applications, investments, and challenges for enterprises. Bus Horiz 58(4):431–440
7. Babu SM, Lakshmi AJ (2015) A study on cloud-based internet of things: CloudIoT. In: 2015 Global conference on communication technologies (GCCT), pp 60–65
8. Parweker P (2011) From the internet of things toward cloud of things. In: International conference on computer and communication technology, pp 329–333
9. Aguzzi S (2013) Definition of a research and innovation policy leveraging cloud computing and IoT combination. Final Report" European Commission SMART
10. Wu M et al (2010) Research on the architecture of Internet of things. Conf Adv Comput Theory Eng 5:484–487
11. Perera C, Georgakopoulos D (2013) Sensing as a service and big data. arXiv Preprint
12. ITU (2012) Overview of the internet of things, pp 22–40
13. Ashton K (2009) That 'internet of things' thing. RFID J pp 49–86
14. Botta A et al (2016) Integration of cloud computing and internet of things: a survey. Futur Gener Comput Syst 56:684–700
15. Alenezi A et al (2017) The impact of cloud forensic readiness on security. In: 7th International conference on cloud computing and services science, pp 1–8
16. Zhou J et al (2017, January) Security and privacy for cloud-based IoT: challenges, countermeasures, and future directions, pp 26–33
17. International Telecommunications Unit (2005) The internet of things. ITU Internet Rep., pp 114–137
18. Joshi GP, Kim SW (2013) Survey, nomenclature, and comparison of reader anti-collision protocols in RFID. IETE Tech Rev 25(5):285–292
19. Suciu G et al (2013) Smart cities built on resilient cloud computing and secure internet of things. In: 19th International conference on control systems and computer science (CSCS), pp 513–518
20. Hartman RD et al (2012) Architecture and measured characteristics of a cloud-based internet of things. In: 2012 International conference on collaboration technologies and systems (CTS), pp 6–12
21. Wang H et al (2014) A vision of IoT: applications, challenges, and opportunities with china perspective. IEEE Internet Things J 1(4):349–359
22. Mohapatra S et al (2010) A survey on applications of the wireless sensor network using cloud computing. Int J Comput Sci Eng Technol, pp 50–55
23. Othman MM, El-Mousa A (2020) Internet of things & cloud computing internet of things as a service approach. In: 2020 11th international conference on information and communication systems (ICICS), pp 318–323. https://doi.org/10.1109/ICICS49469.2020.239503
24. Shanbhag R, Shankarmani R (2015) Architecture for the internet of things to minimize human intervention. In: 2015 International conference on advances in computing, communications and informatics, pp 2348–2353
25. Sharma N et al (2012) Cloud computing for internet of things & sensing-based applications. In: 2012 Sixth international conference on sensing technology (ICST), Kolkata, pp 374–380

26. Zhou J et al (2013) CloudThings: a common architecture for integrating the internet of things with cloud computing. In: Proceedings of the 2013 IEEE 17th International conference on computer supported cooperative work in design (CSCWD), pp 651–657
27. Minh Dang L et al (2019) A survey on internet of things and cloud computing for healthcare. MDPI Electron J
28. Maglogiannis L et al (2012) Bringing IoT and cloud computing towards pervasive healthcare. In: Proceedings of the 6th International conference on innovative mobile and internet services in ubiquitous computing. IMIS, pp 922–926
29. Dar KS et al (2016) Enhancing dependability of cloud-based IoT services through virtualization. In: 2016 IEEE first international conference on internet-of-things design and implementation (IoTDI), pp 106–116
30. Atlam H et al (2017) An overview of risk estimation techniques in risk-based access control for the internet of things. In: 2nd International conference on internet of things, big data and security, pp 1–8
31. Díaz M et al (2016) State-of-the-art, challenges, and open issues in the integration of Internet of things and cloud computing. J Network Comput Appl pp 99–117

Patient Privacy: A Secure Medical Care by Collection, Preservation, and Secure Utilization of Medicinal e-Records Based on IoMT

Haitham Abbas Khalaf

Abstract With the advancement of the Internet of medical things (IoMT) technology, security and privacy have become more important, as this technology allows access to medical equipment or resources from anywhere and at any time. Various methods have been created to provide security and privacy to IoMT, but in this paper, we offer a data gathering methodology that leverages the one-time-pad and symmetric key to protect our data and safeguard the privacy and usage of medicinal e-Records. This technique reduces computing time, increases processing speed, and has a lower implementation cost. The suggested technique is implemented in widely-utilized Python technology, which yields superior findings. Finally, the suggested technique increases security and protects our medical data.

Keywords Internet of medical things (IoMT) · Electronic records (e-records) · Medical data collections

1 Introduction

According to the Internet of medical things (IoMT) [1], billions and trillions of sentient creatures around us will be connected to each other and to their physical surroundings. Based on the progressive accumulation of medical data in a temporarily colonized environment with intelligent items, IoMT systems will deliver a whole new set of sophisticated functions. A healthcare system, a medical data management system, a medical data gathering service, or medical sensors are examples of these. An extensive gathering of medical data, as well as its processing and distribution in the midst of patients' daily lives, raises serious privacy concerns, particularly in areas involving patient privacy, such as wrong diagnosis or failure of treatment [2]. Ignoring these issues may result in unforeseen consequences such as the inability to accept these services, a tarnished image, or expensive legal actions. Privacy is a hot study topic in a variety of technologies and application sectors, with the goal of easing

H. Abbas Khalaf (✉)
University of Anbar, Ramadi, Iraq
e-mail: med.haitham.alakidi@uoanbar.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_21

Fig. 1 IoMT architecture [6]

future IoMT implementation. RFID, WSN, Web customisation, and mobile apps and programs are all examples of this [3]. Even with major contributions from various groups, there is no comprehensive understanding of the expanding IoMT-related privacy problems, since it is a new notion that embraces evolving medical technology and exhibits a variety of dynamic aspects. The rapid growth in the number of smart items (medical equipment) and sophisticated techniques of system interaction and user (patient) reaction are notable among these [4]. These current IoMT features, however, aggravate privacy difficulties and provide unforeseen hazards, necessitating medical technological challenges. These dangers should be considered in light of I the IoMT's reference standards, which define its specific entities and medical data flow, (ii) current medical privacy regulations, and (iii) the IoMT's distinctive and developing aspects [5]. The functioning of newly produced services and the privacy of their users (patients) may be jeopardized due to a lack of understanding of developing difficulties and necessary precautions. The following is how the remainder of the paper is organized: The effort done by several researchers toward privacy-preserving in IoMT is briefly discussed in Part 2. The suggested effort and architecture for implementing privacy in IoMT are offered in Part 3. The experimental data and their analyzing are shown in Part 4. The final part summarizes the work's general conclusion as well as its future prospects (Fig. 1).

2 Related Literature

This part of the study article goes through the previous work done by many academics and authors to build privacy-preserving IoT, as mentioned below: Recently, a number of studies have been undertaken on the stability and privacy of the Internet of things. Defense is the most significant consideration [7], followed by privacy rights [8]. Carrying your own computer is another crucial factor [9] to consider in order to ensure continuing education and e-health situations. Technologies of the digital age (such as big data, the IoMT, the personal medical record (PMR), risk management, computing at a high level, and the cloud) create additional possibilities for healthcare system transformation. Linked medical devices, also known as IoTM, are expected to improve patient safety and quality in healthcare [10, 11]. Despite this, cybersecurity has emerged as a major competitive issue for healthcare providers [12]. When they are classified as susceptible targets with outdated security, poor information systems, and IT regulation, hackers do not hesitate to access hospital databases, paralyze networks with malware, sell patient information to the highest bidder, and threaten to disclose. Furthermore, in present e-health, IoMT has the ability to widen the attack surface. Johnson & Johnson recently admitted that their wireless insulin pumps are vulnerable to hacking [13]. For long time, an independent security professional studied the networking interfaces of the systems while acting as a patient. Despite the little danger of abuse, comparable gadgets are becoming more common in current healthcare technologies (such as defibrillators and pacemakers). This equipment is exposed to a new type of hazard. As a result, risk analyzing is necessary for the healthcare system to have a better knowledge of the security and privacy features offered. As a result, concerns of security and privacy constitute a significant barrier to participation in and hence success in the social health ecosystem. Today, it is apparent that protection is necessary to fulfill all present standards and laws while retaining the flexibility to react to future needs, regulatory requirements, and technical obstacles. The most current solutions in this topic are investigated in this study. It is a step-by-step approach to creating new IoT and IoMT applications while also considering the genetic implications of CE in the healthcare industry. The research focuses on security methods that may be accessible from the system to the cloud (E2E), as well as data collection, transport, and storage, and reusing or discarding relevant equipment. With respect to SGs based on fog computing (FCSG), Feyza Yildirim Okay et al. presented two types of SDA protocol I FCSG-DF (ii) FCSG-P) with efficient and lightweight characteristics, and these two offered a solution for addressing the aforementioned demands. These two SDA methods have a lower processing overhead and a substantially faster reaction time. Furthermore, through medical data aggregation processes that protect patients' privacy, the hierarchical fog computing structure of "FCSG" ensures the privacy of specifically energy-consuming data. The suggested protocols outperform methods supported by the cloud that do not entail data aggregation when it comes to medical data transmission and storage, according to the findings of rigorous performance assessment and analysis. Furthermore, privacy research confirmed that these two protocols effectively protect each layer's data privacy [14]. Mengyao Zheng et al. are a categorization of existing cloud computing-based privacy-preserving machine learning algorithms. It also presents a private information inference strategy in which IoT devices disguise data before transmission utilizing a lightweight neural network and then categorize the obfuscated data with a deep neural network on the cloud. The "MNIST dataset" [15] is utilized to evaluate performance. Prem Prakash Jayaraman et al. address the privacy dilemma by proposing unique privacy-preserving methodologies for IoT data and a privacy-preserving IoT architecture. Nevertheless, by utilizing this approach, it is implemented as an effective proof-of-concept system for data privacy protection. These methods make utilize of many IoT cloud data repositories to keep acquired IoT data private. The design and execution are based on the OpenIoT extension, an open-source IoT cloud solution. The efficiency and performance of experimental assessments and their findings are also mentioned [16]. Yuwen Pu et al. suggested two distinct data aggregation strategies for protecting customers' personal data. The first technique arbitrarily splits the data of the IoT device, preserving just one piece. Then, to communicate other portions to all of the group's devices, symmetric encryption is utilized. Following that, those chopped bits were received, and that single component was combined. The aggregator is then informed of the immediate finding following calculation. To provide secure communication, homomorphic and AES encryption are used. The second plan uses the same slicing method as the first. To prevent the revelation of shared actual device data during mutual data fusion, noise data are used. The aggregator and devices communicate securely using AES encryption. Both of these methods, according to the experimental study, ensure the integrity and confidentiality of IoT device data, hence assisting in the defense against external, internal, and collaborating attacks, and similarly [17]. For fog computing improved IoT, Rongxing Lu et al. suggested LPDA, a lightweight technique for data aggregation privacy preservation. It is based on the One-way Hash Chain, Chinese Remainder Theorem, and Homomorphic Paillier encryption to solve the aforementioned problem and to filter inserted bogus data sooner at the network edge. The in-depth security study performed by LPDA identifies it as a unique privacy-preserving approach that is more secure and improves privacy. Furthermore, rigorous performance measurements have revealed that LPDA is actually lightweight in this sort of IoT [18]. Inayat Ali et al. compared and examined the most recent PPDA procedures. The privacy level, energy utilization by sensor nodes, computing cost, sensor node life, communication overhead, and resilience to malicious aggregator are all elements that are included in this comparative study. In order to offer a full overview of each phase of these operations, the most modern methodologies are reviewed here. They proved that their survey article contains the most up-to-date and comprehensive examination of PPDA approaches [19]. Chunqiang Hu et al. presented a data aggregation strategy for the preservation of consumer private data. This method cuts the data from the IoT device arbitrarily, preserving only a single piece. Then, to communicate other portions to all of the group devices, symmetric encryption is utilized. Following that, the chopped portions were received, and that single component was combined. The aggregator is then informed of the immediate outcome of the calculation. According to the analysis, both data integrity and secrecy may be attained here, allowing it to withstand external, internal, conspiring attacks, and so on [20–22].

3 Suggested Methodology

In this section, we have briefly discussed how the suggested architecture works to protect IoT devices and keep their data private or confidential. We assumed that every IoMT device had a given amount of medical record storage and processing power for these records. A group of IoMT devices belonging to the same medical domain (medical centers, hospitals, etc.), for example, is established. These gadgets all share a secret key that only the user and the aggregator know about. The medical data transferred from IoMT devices are preserved by our approach, which utilize a new architecture. The suggested architecture for privacy-preserving medical data aggregation is depicted in this data flow diagram. It first receives data from multiple nodes or IoMT devices as input and then utilizes a multilayer encryption mechanism to safeguard the data. However, if all of the nodes begin transmitting at the same time, the network's traffic will grow, and the network's life will be reduced, thus we will need an aggregator that is linked to each node to solve this problem. The aggregator takes signals from all nodes and utilizes an aggregation technique to combine them into a single signal before sending it to the sink (receiver end) (Fig. 2).

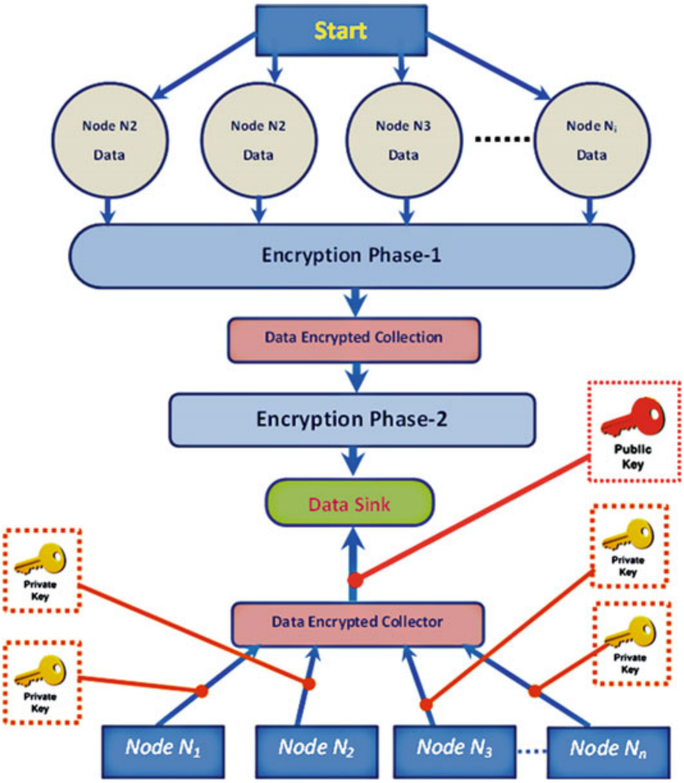


Fig. 2 General outline of research methodology for medical data privacy-preserving

3.1 Architecture

The following is an example of the architecture utilized to create such a system. The suggested system design has the following advantages: 1. simple encryption; 2. simple implementation; 3. quickness; and 4. cost-effectiveness.

3.1.1 Description of Architecture

Let us try to summarize all of the elements, as seen on the previous page's architecture.

1. **Nodes**: When dealing with IoMT, we have a lot of sensors for devices that provide data to the main system. As a result, the system's nodes are all those end devices.
2. **Key**: When encryption is performed, a key is a unique code that is utilized to encode plain text. The cipher text's security is determined by the key's uniqueness.
3. **Encryption**: Encryption is the process of transforming plain information into cipher or encrypted text. This approach turns a human-readable message into a format that cannot be read by humans.
4. **The Collector (Aggregator)**: When all nodes submit data to the main system at the same time, this traffic generates a variety of problems, including increased main system traffic, decreased node life, and network traffic. To avoid such issues, we require an aggregator that is connected to all nodes and collects all node signals before transmitting them to the main system utilizing a data aggregation approach.
5. **Collector (Aggregator) Key**: This is the same as the key, except that it is the unique key of the aggregator.
6. **Sink**: The sink, also known as the main system, is the brain of IoT devices, to which all nodes must submit their data. As a result, it is the system's processing unit.

3.2 The Suggested Architecture Algorithm

The following is a description of the architecture's suggested algorithm:

(1). AGG (EDNi, AgKey); (2). EDNi (EKNi, PTNi); (3). set PTNi to input; (4). UE (ekui, ptui); (5). set ptui to input; (6). utilize calculation matrix; (7). find dkui; (8). find ctui utilizing dkui; (9). return to step ctui (4); (10). repeat for step (2); (11). for all ui $\in$ Ni; (12). repeat step (2); (13). for all Ni;(14). return EDNi; (15). set EDNi to input; (16). repeat step (4); (17). return AEDi; (18). DAED (AEDi, DKi); (19). UD (aedui, dkui); (20). set aedui to input; (21). find option utilizing decoys; (22). repeat step (19); (23). for all ui $\in$ AEDi; (24). return aptui; (25). return APTi.

1. AGG: The following is a description of the architecture's suggested algorithm.

2. Ni is a collection of nodes with i = 1, 2, 3…n.
3. ui is a collection of units in one of Ni's nodes, i.e., ui Ni.
4. EDNi is a set of encrypted data originating from Ni nodes and containing a set of cipher text generated by unit encryption (ctui), i.e., ctui EDNi.
5. EKNi refers to a collection of encrypted keys for Ni nodes that includes a set of encrypted keys for unit encryption (ekui), i.e., ekui EKNi.
6. PTNi is a set of plain text input for Ni nodes that contain plain text that has been decrypted utilizing unit encryption (ptui), i.e., ptui PTNi.
7. The aggregated encrypted data provided by the aggregator from Ni nodes to the sink are referred to as AEDi.
8. The unit encryption (UE) stands for unit encryption.
9. AgKey is the same as EKNi when it comes to the aggregator key.
10. The decryption of aggregated encrypted data is known as DAED.
11. DKi is a collection of dynamic keys utilized in the decryption process that includes a set of dynamic keys for each unit decryption (dkui), i.e., dkui DKi.
12. The unit decoding is referred to as UD.
13. aedui is the collection of encrypted unit data for the ui set of units, where I = 1, 2, 3…n.
14. aptui APTi is a set of resulting plain text that contains a set of resulted plain text for unit decryption (aptui).

The suggested method utilizes a novel approach to data aggregation. As each node requests data, it must be provided to the aggregator, which will aggregate the data and send it to the sink. However, as shown in the design, all data will be encrypted and transferred via each node utilizing its private key. Then, once each node's key and cipher text are combined and sent to the aggregator, the aggregator aggregates all of the data utilizing an aggregation technique and then transmits the data to an encryption method with its own unique key, resulting in multilayer encryption. The encryption modules are all identical. So, first and foremost, we must grasp the encryption module in order to comprehend the entire operation.

Stage 1: Encryption

Let us have a look at the encryption method. We utilize layered encryption, as shown in the above design, to ensure that data security is maximized. The encryption procedure is the same as it has always been:

(1). Key, (2). plain text or data text, (3). algorithm or technique, and (4). cipher text.

Take a look at the unit module of encryption to get a better grasp of the encryption module. Let us take a look at each phrase individually:

1. Encryption: Encryption is the process of transforming plain text into cipher or encrypted text.
2. Plain Text: Plain text is the data or message that has to be translated or encrypted into cipher text.
3. Cipher text: Cipher text refers to the data that is generated as a finding of encryption, or the transformed ex or message.

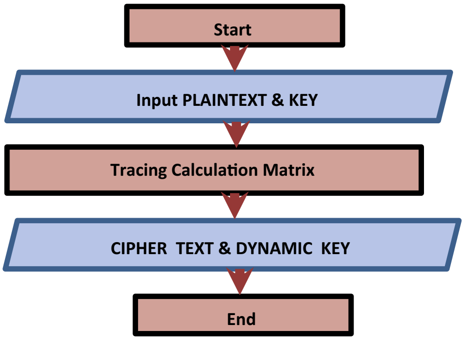


Fig. 3 Unit encryption process

4. Key: The key is the one-of-a-kind code that is utilized to encrypt plain text. The encryption's security is determined by the key's uniqueness.

Stage 2: Encryption-Unit

Because the scheme we are employing is based on a legitimate system or has its own architecture for plain text encryption. Let us talk about it (Fig. 3).

Step 1: In order to build this architecture, we will need two items.

1. The Plain Text 2. TPrivate Key

Step 2: Our suggested approach now relies entirely on a matrix known as the calculation matrix. Now, we need to discover the dynamic key utilizing that matrix. It is a one-of-a-kind key created with plain text and the static key to improve security. So, we have got our calculation matrix.

Step 3: Make an equivalent number out of the message and the key.

Step 4: Calculate the dynamic key utilizing the calculation matrix, with the formula dynamic key (x) = calculation matrix [key(x)] [Text(x)] [Plain text(x)] [Plain text(x)]. We will trace each character of plain text and key for the dynamic key.

Step 5: Now we will utilize a dynamic key to encrypt the plain text. Cipher text (x) = (Plain Text (x) + Dynamic Key (x)) mod 27 is the formula for doing so.

Step 6: Transforming the data into text.

Stage 3: Encryption Findings

Stage 4: Decryption

Let us have a look at the decryption method. We utilize layered encryption, as shown in the above design, to ensure that data security is maximized. The decryption procedure is the same every time: dynamic key, (2) data or cipher text, (3) technique or algorithm, (4) plain text take a look at the unit module of encryption for a better understanding of the decryption module.

Stage 5: Decryption-Unit

As the scheme we are utilizing follows a proper system of decryption so it also has an algorithm for the same.

3.3 Advantages of the Suggested Method

The suggested system has the following benefits over existing approaches.

1. The suggested encryption approach is a free length encryption technique that may be employed as a block encryption technique after a length of 676.
2. Because the suggested encryption method employs symmetric encryption, encrypting plain text is as straightforward as decrypting cipher text.
3. Decrypting the cipher text will take 5×1018 years if an attacker tries to assault privacy by trial and error.
4. Because the suggested encryption approach is simple to implement, encryption and decryption require less time.
5. All of the above approaches rely on one-time data encryption, however, because of the system's design, multilayer encryption may be utilized to increase data security.
6. When compared to all prior suggested aggregation methods, the suggested method is more secure.
7. The suggested encryption approach makes advantage of low-level processing, making it simple to implement.
8. The suggested system implementation is cost effective since it requires little computing labor, requiring fewer resources to implement.
9. The suggested encryption technology is simple to utilize and implement while still providing strong security.
10. Because of the usage of multi-level encryption and multilayer data for encryption, the suggested encryption approaches give the greatest number of conceivable plain text scenarios when compared to all the previous data.
11. Because the suggested encryption algorithms may operate with any length of key, the key can be repeated in a queue for large text.
12. The suggested solution is also appropriate for systems with limited resources, such as those with a low budget, high-security concerns, low compute power, and processing constraints.
13. Utilizing the given architecture, we can easily implement recursive encryption in a system where we do not have any constraints on computation, time, cost, or processing.

3.4 Comparison

The suggested encryption approach will be compared to all prior encryption strategies available at the time.

4 Experimental Findings

Security, computational difficulty, implementation complexity, design complexity, memory consumption, and processing speed are utilized to compare the suggested approach to other current methods such as RSA, DES, and AES. The value of the suggested and existing approach (A: average; V.H: very high; V.L: very low; H: high; L: low; E: easy; D: difficult; A.S: any-size; S: symmetric; S.V: symmetric and vigenere; C.H: comparatively-highest) was employed in this study (Table 1).

4.1 Analyzing of Security

In this area, a security parameter was compared between a suggested and existing approach (Table 2), and a comparison analysis was performed (Fig. 1). We discovered that our suggested approach is much more secure than any other present security method after examining security parameters.

Table 1 Present methods compared with the suggestion technique

	RSA	DES	AES	Suggested aggregation system
Time	V.H	H	A	A
Memory consumption	V.H	H	A	L
Computation need	H	H	A	L
Processing need	H	H	A	L
Input length	128 bits	128 bits	128 bits	A.S
Key length	128/192/256	128/192/256	128/192/256	A.S
Design complexity	D	D	D	E
Implementation	D	D	D	E
Cost	H	H	A	V.L
Security	H	H	A	C.H
Cipher type	S	S	S	S.V
Possible combination	2^{128}	2^{128}	2^{128}	$2^{512} \times 2^{(256 * 4)}$
Time to crack all possible keys	1.02×10^{18} years	1.02×10^{18} years	1.02×10^{18} years	4.08×10^{18} years

Table 2 Security comparison of suggested and existing method

S. No.	Models	Security
1	RSA	High
2	DES	High
3	AES	Average
4	Suggested	Very high

Table 3 ComCos comparison of suggested and existing method

S. No.	Models	Cost
1	RSA	H
2	DES	H
3	AES	A
4	Suggested	V.L

4.2 Analyzing of Computational Cost (ComCos)

In this part, we compare the computation cost parameter of a suggested and current technique, as shown in Table 3, and compare the findings in Fig. 2. We will disregard the cost of hash functions and addition operations in favor of focusing entirely on the cost of encoding and decoding operations because the ComCos of exponential and multiplication operations is significantly more than that of hash functions and addition operations. After researching the ComCos parameter, we discovered that our suggested strategy has a much lower ComCos than the other current security options.

4.3 Analyzing of Implementation Complexity

Table 4 compares the implementation difficulty characteristics of a suggested and current technique, and Fig. 3 depicts the comparative study. According to an examination of the implementation complexity parameter, our suggested approach has a lower implementation difficulty than any other current security technique.

Table 4 Implementation complexity comparison of suggested and existing method

S. No.	Models	Implementation complexity
1	RSA	H
2	DES	H
3	AES	H
4	Suggested	L

Table 5 Design complexity comparison of suggested and existing method

S. No.	Models	Design complexity
1	RSA	H
2	DES	H
3	AES	H
4	Suggested	L

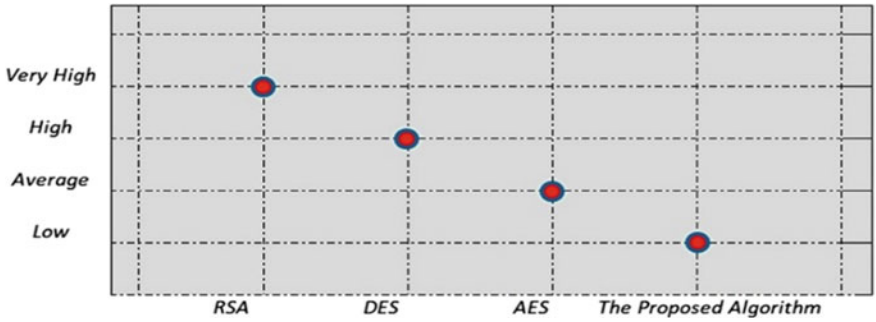


Fig. 4 Analysis of computational requirement

4.4 Analyzing of Design Complexity

In this section, we compare the design complexity parameter between a suggested and existing approach, as indicated in Table 5, and compare the findings in Fig. 4. Our suggested system has a lower design complexity than the other current security solutions, according to an analysis of the design complexity parameter.

4.5 Processing Need Analyzing

In this section, we compare the processing required parameter of a suggested and existing approach utilizing Table 6. According to an analysis of the processing demand parameter, our suggested solution has a lower processing need than the other current security techniques.

Table 6 Processing need a comparison of suggested and existing method

S. No.	Models	Processing need
1	RSA	H
2	DES	H
3	AES	A
4	Suggested	L

Table 7 shows a comparison of the suggested and current methods in terms of computational speed

S. No.	Models	Computation need
1	RSA	H
2	DES	H
3	AES	A
4	Suggested	L

Table 8 Memory consumption comparison of suggested and existing method

S. No.	Models	Memory consumption
1	RSA	V.H
2	DES	H
3	AES	A
4	Suggested	L

4.6 Computational Need Analysis

In this section, we compare the computing requirement parameter of a suggested and existing approach, as indicated in Table 7, and compare the findings in Fig. 6. We determined that our suggested approach takes less computer power than other current security solutions after looking at the computational demand parameter.

4.7 Memory Consumption Analysis

In this section, we compare the memory consumption parameters of a suggested and existing approach, as stated in Table 8, and a comparison analysis is displayed in Fig. 7. After looking into the memory usage parameter, we discovered that our suggested solution utilizes less memory than existing security solutions (Fig. 4).

4.8 Time Consumption Analyzing

Table 8 shows the comparison of time consumption parameters between the suggested and existing methods, the comparative analysis. After examining the time consumption parameter, it was discovered that our suggested solution consumes less time on average than the other security methods already in utilize.

5 Conclusion

(IoMT) is transforming our daily lives by bridging the gap between the physical and digital worlds/Internet. To enable real-time processing, smooth connectivity, and ubiquitous sensing in IoT applications, a privacy-preserving data aggregation approach that utilizes storage and computational resources at network edges is regarded a core ingredient. Various security schemes have been developed to give security and privacy to IoT devices, but in this work, we provide a revolutionary security architecture that is both quicker and less expensive than existing security schemes. The suggested architecture is compared to the existing security methods RSA, DES, and AES utilizing various characteristics such as computation cost, implementation complexity, memory consumption, design complexity, processing demands, computation need, ComCos, and so on. After examining various security techniques, we discovered that our suggested solution has a lower ComCos, utilize less memory, and has a lower design and implementation complexity. Similarly, the suggested technique requires extremely little processing and computing, which improves the efficiency of IoMT devices. This suggested solution improves the security of IoMT devices by combining multilayer encryption with a multilayer of data. Multilayer encryption protects the network against internal, external, and collusion assaults, ensuring the confidentiality and integrity of user data. The suggested technology generates recursive encryption, considerably improving the security of data exchange through networks and IoT devices.

References

1. Evans D (2011) The internet of things—how the next evolution of the internet is changing everything. CISCO white paper 2011
2. David K, Jefferies N (2012) Wireless visions: a look to the future by the fellows of the wwrf. Veh Technol Magaz IEEE 7(4):26–36. doi:https://doi.org/10.1109/MVT.2012.2218433
3. Mattern F, Floerkemeier C (2010) From active data management to event-based systems and more. Springer
4. Presser M, Krco Sa (2010) IOT-I: internet of things initiative: public deliverables—D2.1: initial report on IoT applications of strategic interest 2010
5. Atzori L, Iera A, Morabito G (2010) The internet of things: a survey. Comput Netw 54(15):2787–2805. doi:https://doi.org/10.1016/j.comnet.2010.05.010
6. Din IU, Almogren A, Guizani M, Zuair M (2019) A decade of internet of things: analysis in the light of healthcare applications. Ieee Access 7:89967–89979
7. Jha NK (2017) Internet-of-medical-things, great lakes symposium on VLSI (GLSVLSI), May 2017, Banff, Alberta, Canada, pp 7–7
8. Islam SMR, Kwak D, Kabir MDH, Hossain M, Kwak K-S (2015) The internet of things for health care: a comprehensive survey. IEEE Access 3(1):678–708
9. Xi W, Ling L (2016) Research on IoT privacy security risks, ICIICII, IEEE, Wuhan, China, 3–4 Dec
10. Abie H, Balasingham I (2012) Risk-based adaptive security for smart IoT in eHealth, BodyNets, 24–26 February, Oslo, Norway, pp 269–275
11. Tank B, Upadhyay H, Patel H (2016) A survey on IoT privacy issues and mitigation techniques, ICTCS, Udaipur, India, 4–5 March. Article 2:1–4

12. Miller KW, Voas J, Hurlburt GF (2012) BYOD: security and privacy considerations. IT Profess 14(5):53–55
13. Stathiakis N, Chronaki CE, Skipenes E, Henriksen E, Charalambus E, Sykianakis A, Vrouchos G, Antonakis N, Tsiknakis M, Orphanoudakis S (2003) Risk assessment of a cardiology eHealth service in HYGEIAnet. In: Computers in cardiology, IEEE, 21–24 Sept, Thessaloniki, Greece, pp 201–204
14. Rockoff JD (2016) J&J warns insulin pump vulnerable to cyber hacking—OneTouch Ping uses unencrypted radio signal. In: The Wall Street Journal, 4 Oct 2016
15. Okay FY et al (2020) Fog computing-based privacy preserving data aggregation protocols. Trans Emerg Tel Tech e3900. wileyonlinelibrary.com/journal/ett©2020 John Wiley & Sons, Ltd. 1 of 23 https://doi.org/10.1002/ett.3900
16. Zheng M et al, Challenges of privacy preserving machine learning in IoT. AIChallengeIoT'19, November 10–13, 2019, New York, NY, USA c 2019 Association for Computing Machinery. ACM ISBN 978-1-4503-7013-4/19/11...$15.00. https://doi.org/10.1145/3363347.3363357
17. Jayaraman PP et al (2017) Privacy preserving internet of things: from privacy techniques to a blueprint architecture and efficient implementation. http://dx.doi.org/https://doi.org/10.1016/j.future.2017.03.0010167-739X/© 2017 Published by Elsevier B.V.
18. Pu Y et al, Two secure privacy-preserving data aggregation schemes for IoT. Hindawi Wireless Commun Mob Comput 3985232:11. https://doi.org/https://doi.org/10.1155/2019/3985232
19. Lu R et al, A lightweight privacy-preserving data aggregation scheme for fog computing-enhanced IoT. Article in IEEE Access Mar 2017. doi: https://doi.org/10.1109/ACCESS.2017.2677520
20. Ali I et al, Privacy-presereving data aggregation in resource-constrained sensor nodes in internet of things: a review. doi: https://doi.org/10.1016/j.fcij.2017.11.004
21. Hu C et al, An efficient privacy-preserving data aggregation scheme for IoT. doi: https://doi.org/10.1007/978-3-319-94268-1_14
22. Lavanya K, J Obaid A, Sumaiya Thaseen I, Abhishek K, Saboo K, Paturkar R (2020) Terrain mapping of LandSat8 images using MNF and classifying soil properties using ensemble modelling. Int J Nonlinear Anal Appl 11(Special Issue):527–541. doi: https://doi.org/10.22075/ijnaa.2020.4750
23. Ziegeldorf JH, Garcia Morchon O, Wehrle K (2014) Privacy in the Internet of things: threats and challenges. Secur Commun Netw 7(12):2728–2742

Intelligent Cloud and IoT-Based Voice-Controlled Car

Saroja Kumar Rout, Bibhuprasad Sahu, Brojo Kishore Mishra, Nalinikanta Routray, and Pradyumna Kumar Mohapatra

Abstract The research work depicts an intelligent cloud and IoT-based voice-activated vehicle that responds to voice commands. Noise and distance handling, on the contrary, will require further progress. The car is controlled using speech instructions that are straightforward such as forward, backward, left, right, and stop. An android application sends these commands to the Bluetooth module. We can monitor the car and process the data with the aid of an intelligent cloud computing device. The Bluetooth device and control unit work together to retain and monitor voice recognition. This vehicle then responds to commands obtained from an Android application, allowing the user to control the vehicle via Bluetooth or voice commands, and monitor it using a real-time GPS tracking device. The microcontroller analyzes and executes this instruction. Image processing is applied in the vehicle to become aware of the shadows and obstructions. This car will function without the use of any hard manpower; simply attach your phone to the device, enter the password, and use

S. Kumar Rout (✉) · B. Sahu
Department of Information Technology, Vardhaman College of Engineering (Autonomous), Hyderabad, Telangana, India
e-mail: rout_sarojkumar@yahoo.co.in

B. Sahu
e-mail: prasadnikhil176@gmail.com

B. Kishore Mishra
Department of Artificial Intelligence and Data Science, Vardhaman College of Engineering (Autonomous), Hyderabad, India
e-mail: brojomishra@gmail.com

N. Routray
Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India
e-mail: nalini.bjbcollege12@gmail.com

P. Kumar Mohapatra
Research Scholar, Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India
e-mail: er_pradyumna@yahoo.co.in

Vedang Institute of Technology, Bhubaneswar, Odisha, India

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_22

it as a voice controller or Bluetooth controller. This research has been limited to short-range; however, using long-range modules could result in long-range communication with the vehicle.

Keywords Cloud computing · IoT · Bluetooth module · Speech recognition · Android application

1 Introduction

Speech is a vital aspect of recognizing commands and transmitting them to the vehicle via a wireless device since this paper is connected to a voice-activated vehicle. "Automatic speech recognition (ASR)" is a term used to describe the process of identifying speech. Communication between it and a laptop and a microcontroller, voice input and interpretation, wireless categorical variables delivery from a microprocessor to another microprocessor, and microprocessor-based motor control are all part of this project. This project is about using a Bluetooth recognition device to power a car using voice commands. When we talk about voice control, the first thing that comes to mind is speech recognition, followed by wireless data transformation. Speech recognition is capable of recognizing the specific terms you want to be recognized. The receiver or microcontroller receives the words or data through wireless transmission. These instructions are sent to the Bluetooth module via the Android application. To record and read voice commands, the control unit is connected to a Bluetooth smartphone. This vehicle then responds to commands obtained from an Android application. The user can control the vehicle via Bluetooth or voice control and monitor it using a real-time GPS tracking device [1, 2]. Robotics is one of the most rapidly developing fields of research. The key explanation for this is that robots provide low-cost labor with high-performance precision. Robots have also been shown to transcend human limits, allowing them to complete tasks previously thought to be difficult for humans. The applications for which robots can be used are endless. As a result, the demand for precise and efficient robot control is growing. Controlling robots can be done in a variety of ways. This research paper will concentrate on one of the hundred techniques, namely Closed-loop systems focused on audio channels, which are among the most simple and efficient methods of management of the robot since it communicates with the devices by speaking. During our study, we came up with a device that allows a basic robotic vehicle capable of driving in all four sides, allowing it to perform the following tasks. It has the power to move forward, reverse, move left and right, and come to a complete stop during any time.

2 Methodology

We use a simple Bluetooth control car by speech recognition system called "Smart Car" to control an intelligent cloud computing and IoT powered car using our voice. This car is designed to provide a platform for developing low-cost Bluetooth and cloud-based/voice control cars that have high computing and sensing capabilities thanks to the smartphone that serves as the control system. This system can be controlled by anyone without difficulty. This car will function without the use of any hard manpower; simply attach your phone to the computer, enter the password, and use it as a voice controller or Bluetooth controller. This device can operate on an electrical power supply, similar to a generator, and can be recharged if the power supply battery is depleted. A GPS monitoring system can assist the user when the car is stolen, allowing for easy location tracking. This paper details the design and construction of a smart car that is being considered as a potential solution to the automation car. This technology is rapidly replacing manual labor around the world. The proposed vehicle automation would integrate technology into real-world solutions. The conventional method of driving is too complicated, such as changing gears, breaking, and so on. Whereas in the robotic environment, the user just needs to use his or her phone to power the vehicle with our application. This technology replacement would save resources while reducing manpower. This app can communicate directly with the recipient. The parameters that the user uses to transfer data to the Bluetooth module wirelessly (receiver). The data will be transferred from the receiver to the microcontroller, which will be controlled by the microcontroller. If an obstacle is detected in front of the device, the sensing sensor will detect the obstacle and stop the device, preventing accidents. When the car is stolen, the GPS tracking device can also assist the user by allowing for simple location tracking.

The Arduino SDK Includes (Interface) is a Java program that works on any platform. It has a code editor that includes syntax highlighting, brace matching, and automatic indentation, as well as the ability to compile and submit with a simple click, you can upload programs on the board. The robot compares the command with the stored software and then sets the command using wireless communication as per voice. These suggested methods would be useful for devices such as assistive robotics for people with disabilities or automotive applications such as work robots. We are utilizing it in slave mode in our project, which means it will accept connections from the android app [3, 4]. It requires a 3.3–6 V supply; thus when connecting the Tx pin of the Arduino board to the Rx pin of the Bluetooth module (the Rx pin only supports 3.3 V), we will need to utilize a voltage divider, else the module will simply crash. The Tx pin does not have this need because the module uses the same power source as the Arduino board. We must make use of the module's default password, which is either "0000" or "1234" [5], to connect the module to the android app. Based on the concept, the module has two motors A and B, as well as one for the Ground pin, the motor's VCC, and a 5 V pin that may be used as an input or output. This is contingent on our power supply, if the voltage level is 12 V. If the supply voltage is >= 12 V, it will produce output. It will need feedback to avoid the burnout

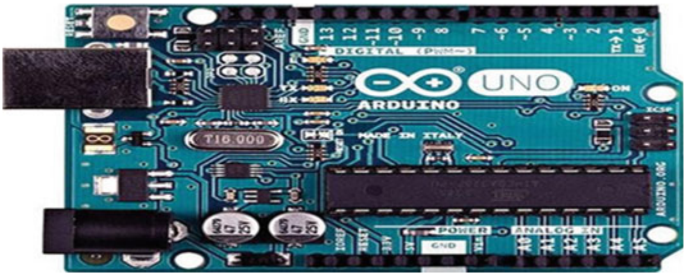


Fig. 1 Arduino Uno device

of other components [6, 7]. The proposed scheme utilizes a smartphone the speech recognition system. We will do this with an Android app that uses Google Speech Recognition to interpret human speech and then uses the Google Speech to Text API to convert it to text [8, 9]. The code that will be communicated to the microcontroller over Bluetooth will be generated from the transformed text. The HC-05 will serve as a receiver (Rx), receiving and transferring data from the smartphone to the decode circuit [10].

2.1 Component Description

2.1.1 Arduino Uno

The Arduino Uno is an ATmega328-based microcontroller board. Among the 20 digital input/output pins are a 16 MHz resonator, a USB connection, a power jack, an ICSP header, and a reset button (of which 6 can be used as PWM outputs and 6 can be used as analog inputs). To get started with the microcontroller, simply connect it to a device via USB or power it with an AC–DC converter or battery. The FTDI USB-to-serial driver chip is not used in the Uno, unlike previous boards. An ATmega16U2 is utilized instead, which has been coded to act as a USB-to-serial converter. A USB boot loader is included with this extra microcontroller, making it simple to develop for advanced users. Arduino is a great place to start if you are new to embedded electronics because it has a large library of support libraries and hardware add-on “shields,” as well as an active support community (e.g., you can easily render your Arduino wireless with our wixel shields). A Spark Entertaining Inventor’s Kit is now available, which includes an Arduino Uno and many components (including a breadboard, sensors, jumper wires, and LEDs) for a range of fun first-time projects. Figure 1 depicts the Arduino Uno.

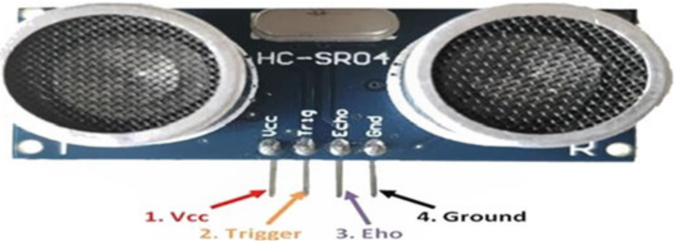


Fig. 2 Ultrasonic sensor

2.1.2 Ultrasonic Sensor

A device that employs ultrasonic sound waves to calculate the distance to an item is known as an ultrasonic sensor. A transducer is used in an ultrasonic sensor to send and receive ultrasonic pulses that offer information about the item's location. Echo patterns are created when high-frequency sound waves bounce off various surfaces. The operation of this module is simple. When it meets an obstacle or object, it emits a 40 kHz ultrasonic pulse that travels through the air and returns to the sensor. By multiplying the travel time by the sound speed, one can calculate the distance. Figure 2 shows how an ultrasonic sensor is used to detect items in this project.

2.1.3 Bluetooth

Bluetooth is a low-power, high-speed wireless technology that allows phones and other portable devices to communicate with each other. It is an IEEE 802.15.1 standard for connecting phones, computers, and other network devices over short distances without the use of wires. Bluetooth is a wireless technology that transmits signals over short distances, normally up to 30 ft (10 m). To achieve this, low-cost transceivers are built into the devices. It operates on the 2.45 GHz frequency band and has a maximum data rate of 721 KBps and three speech channels. This frequency range has been authorized for the use of industrial, research, and medical devices (ISM) by international agreement compatible with 1.0 devices. The Bluetooth (HC-05) module is used as the receiver in this project, which means that any order placed into the Android app is wirelessly transferred to the Bluetooth module (see Fig. 3). Figure 4 described a Block diagram of a cloud computing-based car.

2.1.4 Motor Driver

Motors in autonomous robots are driven by an integrated circuit chip called a motor driver IC. Motor driver integrated circuits (ICs) connect robot microprocessors to the robot's motors. The L293 series, which includes the L293D, L293NE, and others, are the most often used motor driver ICs. These integrated circuits are designed to

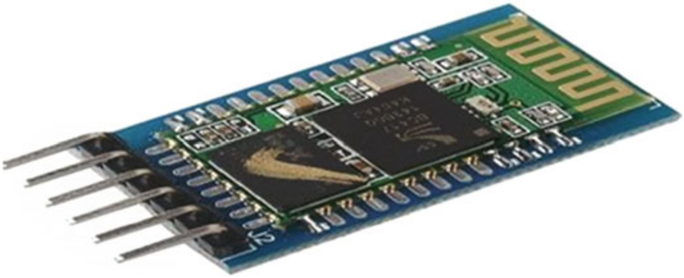


Fig. 3 Bluetooth device

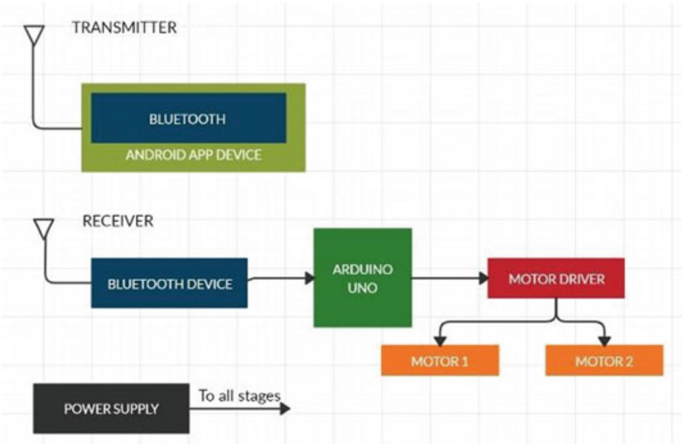


Fig. 4 Block diagram of cloud computing-based car

simultaneously control two DC motors. Just self-driving robots use Motor Driver ICs. Furthermore, most microprocessors operate at low voltages and draw little current, while motors require higher voltages and current. The microprocessor is unable to supply current to the motors as a result. Furthermore, most microprocessors operate at low voltages and draw little current, while motors require higher voltages and current. The microprocessor is unable to supply current to the motors as a result. The microprocessor is unable to supply current to the motors as a result. This is the primary feature of the motor driver IC as described in Fig. 5.

2.1.5 Battery

A chemical process in a battery transfers chemical energy to electrical energy. The chemicals are usually maintained inside the battery. It is utilized to keep track of other circuit elements. Direct current (DC) electricity is generated by a battery (electricity

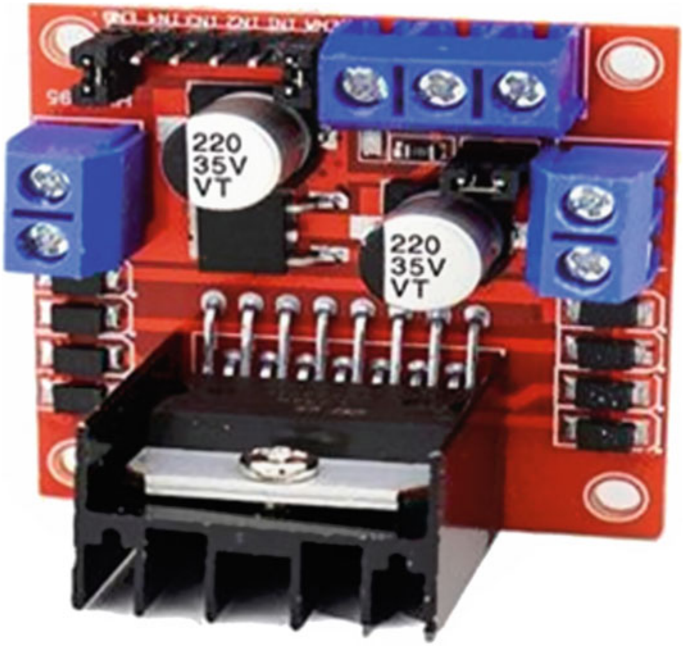


Fig. 5 Motor driver ICs

Fig. 6 Battery

that flows in one direction and does not switch back and forth). Although it is easier and more efficient to use energy from a building outlet, a battery can offer electricity in locations where there is no electric power distribution. Figure 6 shows how it can be used to track moving objects like electric vehicles and telephones.

Fig. 7 Wheel

2.1.6 Wheel

For locomotion, God created legs, and man invented wheels, which is considered one of the greatest inventions of the human period. Wheels are the perfect option for robots because they are simple to build, implement, and use in robots that need to move quickly. They also do not need complex models, prototypes, or algorithms because the robot's center of gravity remains constant regardless of whether it is moving or stationary. They have the drawback of being unstable on rough or rocky terrain, as well as on highly smooth surfaces, where they appear to slip and skid as described in Fig. 7.

2.1.7 DC Motor

Almost single mechanical advancement we see is the result of an electric motor. An electric machine is a sort of energy converter. Electrical energy is converted into mechanical energy by motors. Hundreds of machines that we use daily are powered by electric motors. Direct current (DC) and alternating current (AC) motors are the two most common types of electric motors. The DC motor and how it works will be discussed in this article. Furthermore, what is the operation of a gear DC motor? A direct current motor is an electrical motor that runs on a direct current (DC motor). The operation of any electric motor is dependent on electromagnetism. When a current-carrying conductor is exposed to an external magnetic field, it is subjected to a force proportional to the current and the external magnetic field's power. It is a machine that converts electrical energy to mechanical energy. It is based on the fact that in a magnetic field, a current-carrying conductor is subjected to a force that causes it to rotate relative to its initial location. Figure 8 depicts the situation.

2.1.8 Jumper Wires

Jumper wires are simple cables with connector pins on both ends that can be used to link two locations without the use of solder. To enable fast circuit adjustments,

Fig. 8 DC motor

Fig. 9 Jumper wires

jumper wires are frequently used with breadboards and other prototyping equipment. It is a really simple process. The most common electrical component is jumper wires. The colors of jumper wires have no meaning, even though they exist in a range of colors. A red jumper wire is the same as a black jumper wire in this case. However, as shown in Fig. 9, you can utilize the colors to distinguish between different sorts of connections, such as ground and electricity.

2.1.9 Node MCU

A low-cost IoT platform is the open-source Node MCU platform. It included firmware that ran on Espressif Systems' ESP8266 Wi-Fi SoC, as well as hardware that was based on the ESP-12 module at the time. The ESP32 32-bit MCU was later added to the list of compatible devices. The ESP-12E module features an ESP8266 chip with a Ten silica Xtensa 32-bit LX106 RISC microprocessor, which is included with the Node MCU ESP8266 development board. This microprocessor runs at 80–160 MHz and supports RTOS. To store data and programs, the Node MCU has 128 KB of RAM and 4 MB of Flash memory. It is suitable for IoT projects because of its high processing speed, built-in Wi-Fi / Bluetooth, and Deep Sleep Operating features. The Node MCU is controlled through a Micro USB jack and a VIN pin (External Supply Pin). As seen in Fig. 10, it features UART, SPI, and I2C interfaces.

Fig. 10 Node MCU

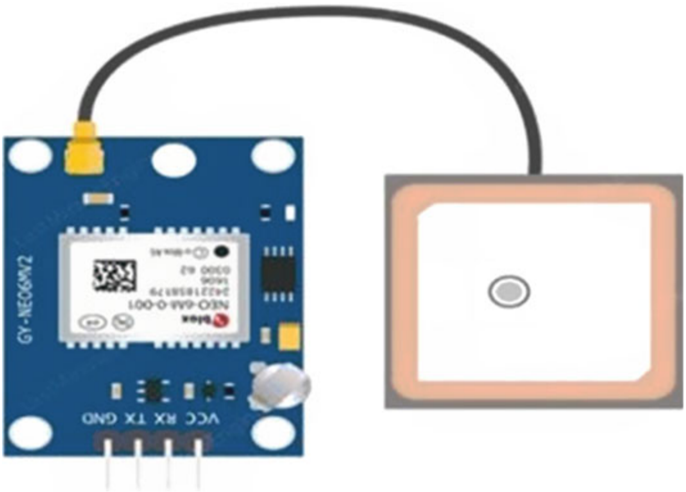


Fig. 11 Neo-6m GPS module

2.1.10 Neo-6m GPS Module

The Neo-6m GPS module is a high-performance GPS module that can be used in a wide range of applications. The power and signal indicators allow you to keep track of the module's status. This board contains a voltage regulator, as well as a logic level converter, allowing it to work with a 3.3–5 V interface. It can bind directly to 5 V Arduino, 3.3 V Arduino Pro Mini, and other similar boards, as well as Node MCU. You can save your setup data, assist now offline data, or receive configuration using an EEPROM and a flash memory chip. GPS-equipped module (3.3–5 V interface with EEPROM and flash) Logic level converter and 3.3–5 V protocol compatibility are included in the UART interface. Antenna is embedded with the GPS module to

Fig. 12 Hotspot device

transmit and receive high gain signal. Figure 11 shows the dimension of the board, which is 25 × 25 mm and 23 mm * 40 mm.

2.1.11 Hotspot Device

A hotspot is a physical site where individuals may connect to the Internet using a wireless Local Area Network (WLAN) and a Router connected to an Internet Service Provider, commonly via Wi-Fi (ISP) [9]. When the Node MCU is connected to the Internet, this system is used to provide Internet access so that we can monitor the live location from anywhere is represented in Fig. 12.

3 Proposed System and Description

This suggested device, we offer a wide range of services which Robotics research on control-type versions. It demonstrates that by utilizing simply voice (human speech) as a control method, real-world artifacts can be investigated and controlled efficiently. This lookup's goal is to design a minimal robotic hardware design that allows this form to concentrate on the infrastructure for Bluetooth communication. It also promotes academic robotics by allowing people to make their own low-cost robots. When the system is running in the application, to do so, a smartphone microphone is employed to recognize commands using your voice. The application translates instructions and converts voice to text inside the app using Google's speech recognition technologies. The text will subsequently be delivered through the receiver part connected through Bluetooth.

The Arduino device has a microcontroller with a lot of features with 32 KB of ISP flash memory, 2 KB RAM, and 1 KB of EEPROM. Serial communication is supported by UART, SPI, and I2C on the panel. The clock speed of the MCU will be 16 MHz. In this version, output pins 3, 4, 5, and 6 digital I/O pins on the Arduino have been programmed as output pins. Arduino pins 0 and 1 are used to make serial communication with the Bluetooth device. The UART serial conversation protocol

Table 1 Features of voice command

Voice command	Function
Forward	Car move in forwarding motion
Backward	Car moves in a backward motion
Right	The car turns right and move
Left	CAR turns left and move
Stop	Car stops, no operation

is used to send text from the Bluetooth device to the Arduino UNO microcontroller panel [11]. Table 1 lists the speech instructions that were used to control the car, as well as their responsibilities. The tracker identified the car location by using a node localization algorithm which is represented in the localization system [12] (Figs. 13 and 14).

An Android device is used to send voice commands to the robotic system through Bluetooth. The robotic system receives these commands through a Bluetooth module that has been mounted. The speed of the vehicle is regulated by the motor driver circuit. A 12 V rechargeable battery linked to the system provides power to the entire electronics.

After the system has been successfully linked, click the Bluetooth textual and iconic push button after opening the app on your phone. Now, you can see the total number of connected devices. To attach the phone to the receiver side, the HC-05

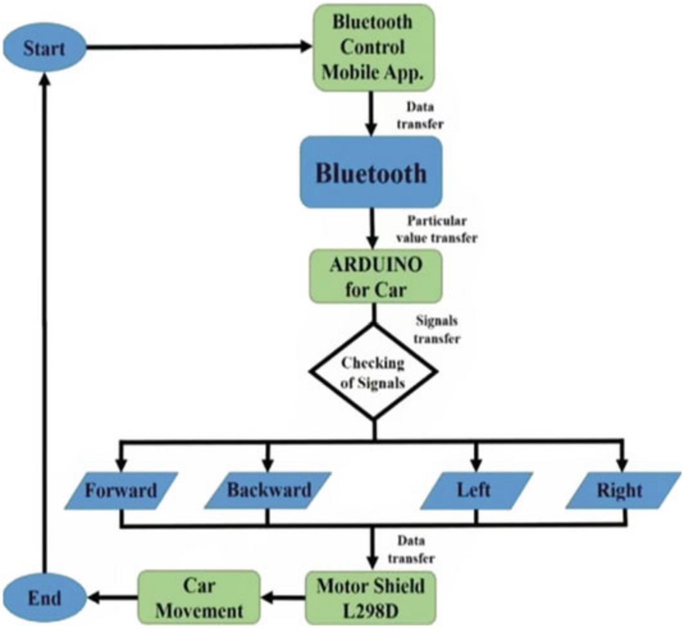


Fig. 13 E-R diagram of intelligent cloud computing and IoT-based voice-controlled car

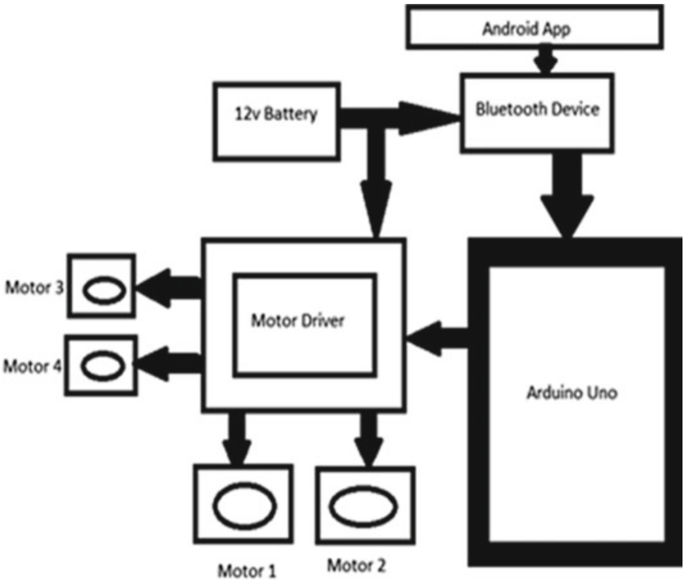


Fig. 14 Block diagram of intelligent cloud computing and IoT-based voice-controlled car system

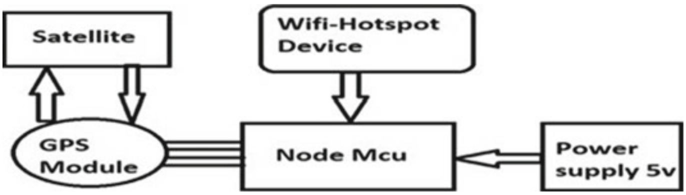


Fig. 15 Block diagram of tracking module of intelligent cloud computing and IoT-based voice-controlled car system

Bluetooth module, select HC-05 from the list. When it appears, the app detects voice commands, converts them to text, and sends it wirelessly through Bluetooth to the receiver. Arduino double-checks the text on the receiving end. The car's movements are governed according to the specification if the string matches. Figure 15 describes the block diagram of the proposed system. The GPS Module and main vehicle circuit diagrams are shown in Figure 16.

4 Performance Analysis and Result Discussion

The voice-activated vehicle is now operational, and the bot takes care of everything. The car is wirelessly connected to the phone, unlike a DTMF robot, allowing the

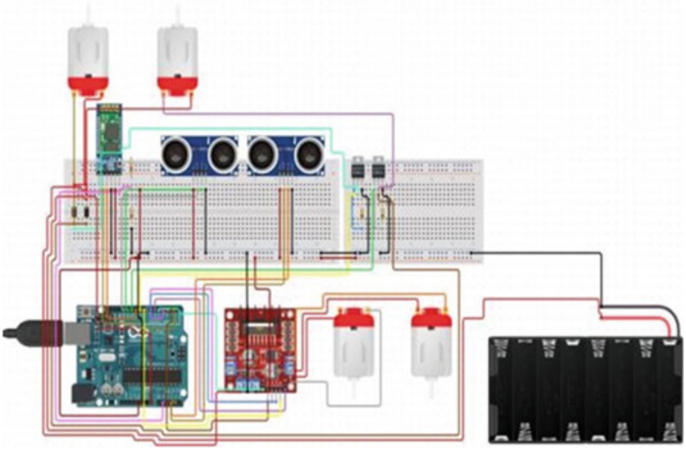


Fig. 16 Circuit diagram of main vehicle module

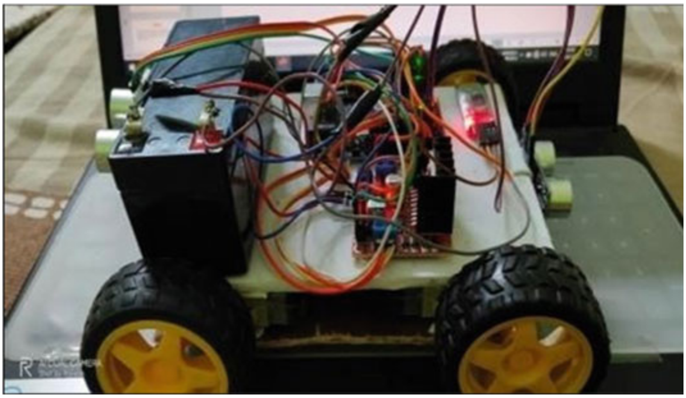


Fig. 17 Proposed voice-controlled

user to keep track of the vehicle. The line follower robot can only travel along a predetermined path; if an obstacle blocks its path, it will remain stationary until the obstacle is removed. Instead of using an offline Bit Voicer server, users can use Google's speech recognition technology to provide voice feedback in a variety of languages. Figure 17 presents the final system model of the proposed system. The cloud-based system is described in Fig. 18.

Voice-controlled car uses the IoT and cloud technology which helps in updating a user with voice command and control the car and also helps them to find the vehicle position by the tracker of this vehicle as represented in Fig. 19. The Bluetooth module is shown in Fig. 20 as being connected through Bluetooth also describes the status connected of the device. The tracking status of the car is described in system to identify the device and location.

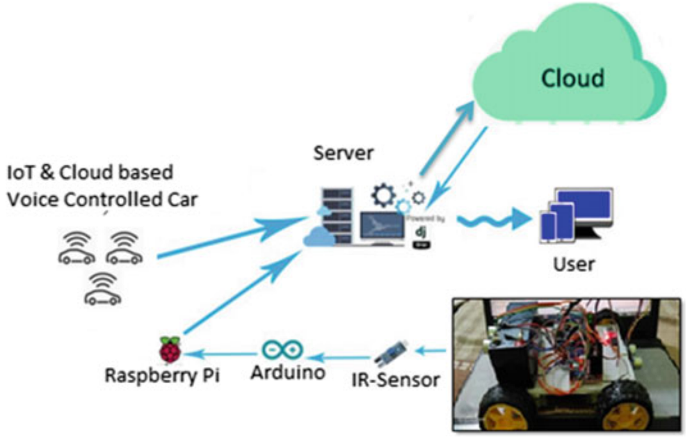


Fig. 18 Voice-controlled car architecture

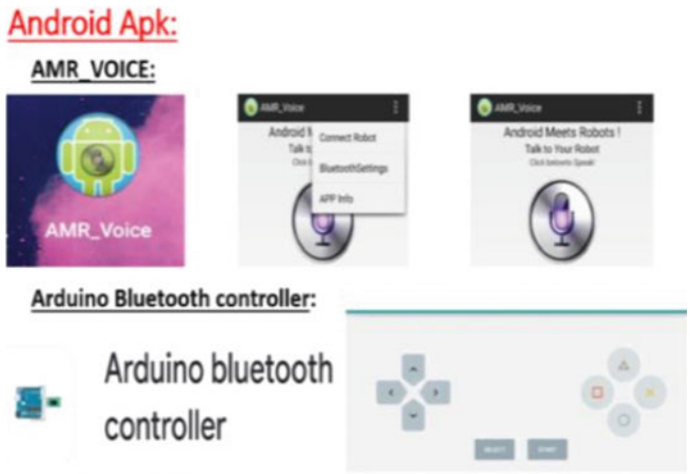


Fig. 19 Status connected

5 Conclusion and Future Scope

This paper outlines the concept of voice control for a home assistant robot. Using a wired network, the server immediately receives the order of speech signals. The car is based on a microcontroller-based architecture for the most part. The original experiments' success is measured, and the results are positive. Potential technical advances include households, schools, vehicle networks, and companies. Noise's impact on speech-to-textual material translation is one of the topics that should be addressed. The robotic behavior is no longer influenced by the speaker's accent, as a cloud-based server interprets speech commands that is unaffected by the accent of the speaker. Energy sources that are renewable for robotics can boost the value of robotic energy while remaining environmentally friendly. Solar cells have the potential to be a viable source of energy. The robotic assistant's architecture is suitable for a variety of tasks, such as chemical processing and at-home relaxation. This experiment was confined to a 100-m-range ZigBee system connected to the car via long-range modules over a long distance. Maintain a consistent sleep and wake-up schedule to boost your energy levels. Image processing can be used in the car to detect obstacles and shade. Servo motors can be used to provide additional proper service. The robot's attention could be tracked by an automated targeting system (Fig. 21).

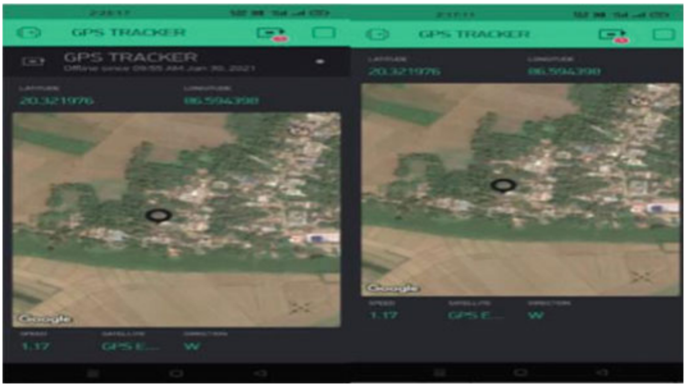


Fig. 20 Identify the device and location

Fig. 21 GPS tracker of this vehicle

References

1. Timsina A, Sharma DC, Sharma N, Bhutia A, Pardhan S (2020) Voice command robotics car. Int Res J Modern Eng Technol Sci 2(06)
2. Gulati U, Dass R (2020) Intelligent car with voice assistance and obstacle detector to aid the disabled. Proc Comp Sci 167:1732–1738
3. Neethu KN (2021) Humanoid robot with surveillance and virtual voice assistance. Turk J Comp Mathe Educ (TURCOMAT) 12(12):2222–2229
4. Santhakumar M, Saran S, Sathishkumar S, Gunaseelan M, Elango K (2021) Voice controlled car using android, arduino and bluetooth. J Elect Eng Technol (IJEET) 12(3):244–251
5. Kanade P, Kanade S (2020) Raspberry pi project-voice controlled robotic assistant for senior citizens. Int Res J Eng Technol (IRJET) 7(10):1044–1049
6. Gupta N (2021) A novel voice controlled robotic vehicle for smart city applications. J Phys Conf Ser 1817(1):012016 (IOP Publishing)
7. Srivastava S, Singh R (2020) Voice controlled robot car using Arduino. Int Res J Eng Technol 7:4033–4037
8. Potts J, Sukittanon S (2012) Exploiting bluetooth on android mobile devices for home security application. In: 2012 Proceedings of IEEE Southeastcon (pp 1–4). IEEE
9. D'Mello S, McCauley L, Markham J (2005) A mechanism for human-robot interaction through informal voice commands. In: ROMAN 2005. IEEE international workshop on robot and human interactive communication. IEEE, pp 184–189
10. Babu BS, Priyadharshini V, Patel P (2021) Review of voice controlled robotic arm-raspberry Pi. Europ J Electr Eng Comp Sci 5(2):1–4
11. https://create.arduino.cc/projecthub/Yug_Ajmera/
12. Rout SK, Rath AK, Rout BR (2016) Efficient energy utilization and node localization in dynamic DV-Hop algorithm for wireless sensor networks. Indian J Sci Technol 9(30)
13. Zhang J, Wang M, Li CY, Li XM, Xu W (2012) Design of intelligent-tracking car based on infrared photoelectric sensor and speech recognition technology. J Chongqing Univ Technol (Nat Sci) 7

Software Testability (Its Benefits, Limitations, and Facilitation)

Jammel Mona

Abstract Software testing refers to a testability method which has test support to improve and predict the software testability. Various types of method have been adopted by researchers and practitioners to improve the testability mechanism in software testing domain. This paper main objective is to reviewing the body of knowledge in this domain and provides a comprehensive overview to new readers and researchers about the software testability. This review selected eighteen papers as evidence to discuss the benefits, limitations, and proposed methods in the domain of software testing. We believe that this short review will give a quick overview to new researchers and readers in the field of software testability.

Keywords Software testing · Issues · Revealing · Facilitation

1 Introduction

Software testing is one of the significant and fundamental aspects to ensure the software system quality. Software testing is the ability to test the software artifact such as requirements, modules, and documents. However, some of the software are not tested easily, and some are easier to test. If the software artifact is high, then the faults findings are tested easily. On the other hand, the lower degree artifact increased the test efforts and has less chances of finding software defects. A better testability of software increases the quality of test activities and decreases cost. Basically, the software testing is based on six factors including software implementation, representation, built-in test capabilities, test cases, processes, and environment. In order to improve and predict the software testability, a various different types of techniques have been designed and proposed by researchers. The huge number of efforts make it difficult to cover all the previous studies in the field of software testing. In addition, we also observed that there is not any clear guidance to measure the testability and address the issues.

J. Mona (✉)
University of Kufa Iraq, Najaf, Iraq
e-mail: jammelm.mana@uokufa.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_23

In this paper, we reviewer some most important survey and review papers on software testing domain. The main objective of this short review is to highlighting the most important and specific aspects of software testability. In order to address the need of software testing, a systematic review conducts related to existing state-of-the-art literature and present the summary of results in the previous studies.

The reminder of this paper is organized as follows: Sect. 2 illustrates the brief overview of software testing. Section 3 presents the existing literature summary and highlighted the main points in the previous literature.

2 Software Testing Overview

The software testing has been offered by IEEE standard, ISO standard, and various other well-known organizations. The software testability is broadly characterized into three groups as shows in Fig. 1.

The first group has definitions, and facilitation related to testing to evaluate the test efficiency of software. The second group is interpreting the testability to reveal the faults related to test effectiveness. The third group is based on other aspects of software testing. The literature indicated that the software testability is classified based on its definition.

Software automation and software disaster testing in many applications have a sequence of activities, processes, and tools. These are used to process in order to sum the test on software. The major software automation activities are test plan, test design or implementation, execution of test, test evaluation, and analysis. Figure 2 presents the testing and automation processes.

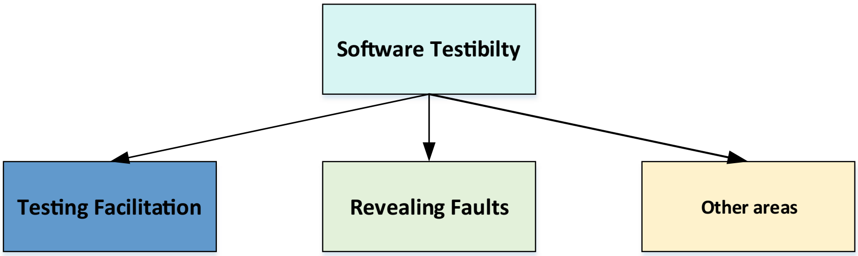


Fig. 1 Main groups of software testability

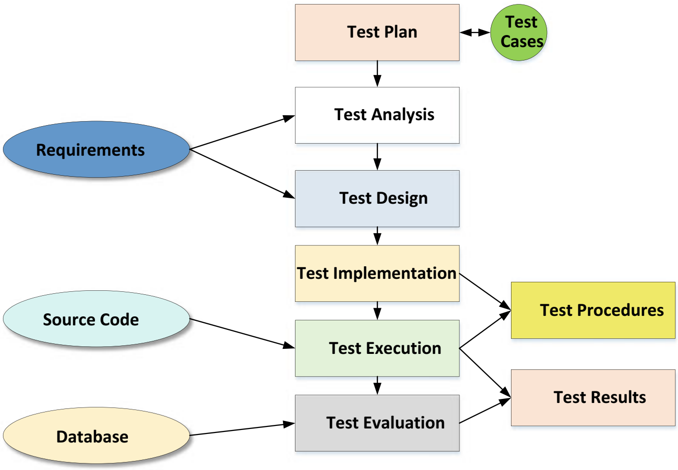


Fig. 2 Testing and automation processes

3 Related Work

Wohlin [1] discussed snowballing approach based on the results of several literature reviews. Snowballing is the principle of sorting through large amounts of data and picking out relevant information. It is the extraction of interesting (non-trivial, implicit, previously unknown, and potentially useful) patterns or knowledge from the huge amount of data. The data in repositories, data sets, Webpages, or databases are mined using snowballing algorithms. These algorithms are divided into three categories: supervised learning algorithms; takes certain labels as inputs (data for weather forecast, fingerprint recognition), which are then used for prediction and modeling, unsupervised learning algorithms; no labels are given as input, and general information is extracted out of the data, and semi-supervised learning algorithms; which use both labeled as well as unlabeled data to perform a certain task. Since snowballing is a very vast domain, it is further divided into sub-domains. These include Web mining, repository mining, big snowballing, spatial data mining, and database mining among others. As snowballing helps to take efficient and effective decisions, which is used in various fields including retail industry, healthcare, insurance, sales/marketing, forensics, fraud detection, intrusion detection, bio-informatics among others. The aim of this paper is to provide a state of the art in systematic literature studies rather than using old database approach, the different snowballing functionalities, and the algorithms which are used on the various kinds of data sets in repositories, databases, and the World Wide Web.

The paper will classify and evaluate the literature based on snowballing functionalities and the type of algorithm used. This will help in a comparative analysis of snowballing techniques, strategies, and algorithms based on their categories and functionalities. It will also help in the identification of algorithms which provide a certain functionality and identify that for a certain category, which is the best strategy that can be implemented, based on the comparison of different approaches.

Garousi et al. [2] discussed and characterized the set of collaborative industry and academic R&D project in the domain of software testing with respect to the existing literature. This study also discussed the percentage of success full industry and academic collaborative projects. Author selected 10 projects (six completed, 2 failures, 2 ongoing) in the domain of software testing. Author highlighted the active role of these projects based on his experience to do work in these projects. These projects had active role to determined and analyzed the existing challenges, patterns, and anti-patterns. This study showed some findings and evidence based on recommendation, for example: It has been observed that the many projects are seem to correct in terms of different aspects but mostly disagreement in confidentially agreement is one of the significant challenges which can cause of project failure. This study recommended that all the parties have considered all the challenges on the proactively and involved with every step in the R&D to reduce the failure in projects. This study encouraged the new software researchers and software engineers to help in their project. The author believed that the software engineer researchers have good future which need better collaboration between the academic and industry with need to interact throughout the process of R&D development.

Shaheen and Du Bousquet [3] discussed software testing and metrics used for testability. As software testing is an expensive activity to be performed on a system so it is very critical that we design such a system which is easy to evaluate and decide its test cases before the actual development of the software system is done. As software testing domain is huge and complex that is why it is difficult to gather all requirements which need to be tested in a software. Many metrics are not usable, and or they are implemented in old tools. The paper has mentioned that there are already 21 metrics presented by Binder in 1994, and they are going to present 40 metrics which can be evaluated on code to increase the code's testability of a software. This paper provides with easy to use metrics for testability of the source code.

Ertam and Aydın [4] discussed machine learning for learning purposes which gives good results specially to find the unknown patterns. Deep learning is subfield of machine learning; it has excellent results as compared to machine learning. Big difference is that; deep learning is when use full when we have large amount of data set. Main difference between machine learning and deep learning is that in machine learning human extracted data for neural network but in case of deep learning, neural network automatically extracts features and pass into hidden layers. In this paper, famous google library TensorFlow has been used. For training purposes, famous MNIST data have been used. In this paper, different activation function has been used with convolutional neural network. As we know that in MNIST data set, large number of images are present. Activation function is basically applying on every neuron output in hidden layer and output layer. These activation functions are used for nonlinear space. In this paper, function used such as rectified linear unit, hyperbolic tangent, exponential linear unit, and sigmoid. At the end for calculating probability of different outputs and select the highest one (output). In MNIST data set, handwritten images are present for best accuracy. Data set save in IDX file format. This is use for representation of vectors and metrics. Data have been divided into training and testing data. For training purposes, data have been feed with labels. In this paper, 100,000 iterations can be carry out. Each picture is feed in 28 * 28 pixel. In convolutional layer, pooling layer and fully connected layer are used. Biases are also use in each neuron output compilation. In last step also known testing data or evaluation step, different activations functions are used. Best result given by ReLU activation function 98.43%. As we increase iteration accuracy increase. If we add different neural network, system may be accuracy increase.

Huda et al. [5] discussed software testability which is very important factor in software development process. Testability of software is always increasing for the software process lifecycle and check the ability of software. Mostly, researchers claimed the testability have major impact on software quality. This report proposing the comprehensive evaluation for testability issues, limitations, and to investigate the less testability factors which are commonly acceptable using systematic review. In this article, firstly, authors have to know the background of testability and its different parameters used by most the practitioners in software development process. Number of Software Testability Survey Reviews according to different software testing factors which are proposed by different researchers. By reviewing different literature, it proves that the time testability is done at the later stages of software development process. So it is easy to say that software testing is done at later stages which leads toward the unstable software it terms of design, use case design. So it is better to measure testability at the early stages of software design to ensure the quality, reduce testing time, and reduce reword for software.

Häser et al. [6] discussed advance systems that are continually interacting with Internet such as cyber-physical systems. In these systems, software and service as software are communicating vary securely. For such system, security and system performance play a vital role. Integration testing is an effective and efficient technique for testing and guarantee the functional requirement, non-functional requirement, and interpretability of a system. To overcome these challenges, test automation and model base testing are good choice for early definition and early validation. These techniques will help us to reduce the upcoming security challenges and issues. The objective of this research paper evolves around model base integration testing approaches. Although a systematic overview and measures for guiding the integration process to test the functional requirements of system can be gain through study of an underlying software. The primary goal is to analyzed the existing stage of model-based integration testing. For synthesize the relevant study, as per guidance of Kitchenham, systematic literature review was conducted in which 713 publication was collected, and then out of these publication, 83 publication is selected for model base integration testing. In addition, the 83 relative publication was extracted for relevant study.

Garousi [7] presented the analysis after doing a survey, based on 289 papers published by the Turkish software engineering research community in the domain of software engineering (SE) from 1992 till 2014. As per their analysis, the author found that the best university is Middle East Technical University, and the top-ranked researcher is Ayşe Başar Bener. Turkey has produced 0.49% of worldwide

SE knowledge, which is quite low and also shows the lack of variety in the general SE field. The positive thing which author found in the analysis is 89 papers (30.8% of the total) shows the International collaboration, namely USA, Canada, and Netherlands, which is good sign. The research papers are mostly based on the academia side only, and the involvement of industry was quite low. Another important finding is the citation to Turkish SE papers in the domain of SE are very poor, which is an alarming situation. The author defined that the Turkish software engineering research community needs to increase the number of publications in the field without compromising on quality. The author has done the analysis, which was more focused on number of publications in SE (quantitative), but the analysis on quality of the publication is missing, which is the future research area. Quality of the papers is also an important factor because without quality, the research work impact is nothing. So finding quality in a publication is an area where research can be done. In this paper, author relates software testability and software performance. Software performance is non-functional property. Software performance measures the system or component functions within given constraints (speed, accuracy, memory). Software performance always deal with time factor which include timeliness and response time.

Hassan et al. [8] presented the software testability where the software is tested from different angles, and the testing efforts are carried out where the software performance is always check timeliness and response time. Software testability is always give the most importance, but software performance is always ignored. Software testability is achieved through different domain with real-time systems. Software testability is divided in few categories like *observability, controllability, automation,* etc. Observability *is around 50%, controllability* is around 46.1%, while *automation* is around 7.7%. Similarly, software performance is divided into response time (23.1), timeliness (46.2%), and memory usage (11.5%).

Suri and Singhani [9] discussed the software testability which has issue like controllability, observability, and testing efforts while for software performance, issues are timeliness and response time. There are many software testing technique and tools which help to reduce the testing efforts and predict the weaknesses of the developed software. The main objective of this paper is to help in finding out which measurement technique is best for software testability at various phases of the life cycle of the software. The main purpose of this research is to test different object-oriented testability metrics from design to development phase. The findings of this research are to provide with an insight of different matrices and defining about the strength and weaknesses of each testability metrics. The outcome of this research is it provides an object-oriented testability metrics which helps in reducing effort in design as well as development phase of software development. During the study, the authors found out that the techniques used in software testability are more in implementation phase as compared to design phase. They concluded that despite all the research and efforts in testability, no one came up with some standard techniques in object-oriented software testability.

Felderer et al. [10] discussed this modern information technology domain, and there are many interconnected networks like cloud computing, social networking, and Internet of things. Security of the sensitive data using these networks always remained an issue as these are always vulnerable to the security threats. In order to improve the security of a particular software, security testing techniques are used to ensure the security and to identify vulnerabilities. Apart from the previous researches, this paper uses the model-based security testing (MBST) techniques using specific classification criteria, i.e., filter and evidence criteria. Filter criteria provide relevant test cases whereas evidence criteria describe the limit of evidence available for usefulness of model-based security testing. Results of over 119 selected research papers are evaluated to check the existing taxonomy versus the new proposed model. An aggregate view and analysis are provided for state-of-the-art model-based security testing.

Petersen et al. [11] presented a systematic reviews which are to gather and compile evidence at one place. Systematic mapping and systematic literature review are some of the approaches used to structure this kind of research. For systematic literature review, many suggestions have been made but for systematic mapping, and the guidelines have not been updated since 2008. That is why author thinks that the way researchers conduct the process of systematic mapping needs to be studied again, and suggestions should be made for the guidelines of systematic mapping. The author tried to find the improvement potential in the guidelines for the systematic mapping. Author also tried to identify how systematic mapping is applied for study selection, analysis, and presentation of data. For this purpose, authors took a systematic mapping study of all those aforementioned tasks following the current guidelines of systematic mapping. The paper concluded that the current guidelines for systematic mapping are not sufficient. Current guidelines only represent few of the activities which actually get conducted in a systematic mapping now a day. The paper also came up with a comprehensive list of activities which now a days are in practice for systematic mapping study. In the end, paper suggested some amendments in the current guideline for conducting systematic mapping. Paper also presented a method to evaluate the quality of mapping study.

Petersen et al. [12] presented the software engineering for programming testing which is been performed. The target that has been characterized in the paper is to direct a methodical mapping study in programming building and give rules by contrasting foundational maps and fundamental surveys and how to pick between these two and how to pick between these two. As this correlation is help in giving the guidelines that are to be set for the efficient mapping. To accomplish this goal, a total orderly procedure was characterized and connected to finish mapping study and a correlation was performed on precise maps with efficient surveys by methodically. The basic procedure ventures of the deliberate mapping study are meaning of research questions, directing the scan for important papers, screening of papers, key wording of digests, and information extraction and mapping. Each procedure steps have a result, the ultimate result of the procedure being the deliberate guide. In the outcomes, a procedure for programming building orderly mapping investigations and contrast it with methodical audits. In perspective on this, rules for doing exact maps are described, and according to results, systematic maps and reviews are assorted similar to targets, extensiveness, authenticity issues, and proposals. In view of this, rules for doing precise maps are characterized, and as per results, systematic maps and audits are diverse as far as objectives, expansiveness, legitimacy issues, and suggestions. In this way, they ought to be utilized correlatively and require distinctive strategies.

Suri and Singhani [13] discussed the software testability which is one of the most important aspects in development and for designing phase of object-oriented testing. By using this testing technique, we can make a few changes in development phase and can prepare new testing techniques. In this research paper, the author is trying to convey that many scientists have declared that testing plays an essential role in development, even then it is not getting used in most software engineering steps. So, the author discusses different techniques, processes, and ways in which object-oriented testing can be implemented on different stages of software development life cycle. Furthermore, the aim of author is to change most people opinion and views on testing and to enhance object-oriented testing techniques and the ways in which these techniques can make a better software.

Binder [14] discussed the software testing which has a direct impact on quality and reliability of software. According to software development, life cycle testing is one of the phases upon which considerable amount of time is consumed. This study focusing on to achieve high reliability in the domain of object-oriented development on the fly. Testability decreases absolute test cost in a unwavering quality driven procedure and increase reliability in a resource limited test process. In this way, notwithstanding of the testing methodology utilized, it is in light of a legitimate concern for an advancement association to improve testability. Testability of an item situated framework is an aftereffect of six essential elements: portrayal, usage, worked in test, the test suite, the test support condition, and procedure capacity. Every impact controllability and discernibleness of the usage, or on the other hand fundamental operational adequacy of the test procedure. About every one of the systems and innovation for accomplishing high testability are settled, be that as it may require budgetary duty, arranging, and cognizant exertion. The progressed worked in test capacities outlined here do not yet exist, yet are possible with existing technology.

Nikfard et al. [15] discussed about the testing issue of designs in software engineering and evaluate different approaches for model testability. The design testing is very important, but it is not implemented before using as a model-based product. Due to this low quality of software, the testing phase is very important to expose the faults and issues of a software. In the past, the testing of model was done often when design was completed. Due to this, we cannot determine the level of model testability of the design efficiently. If testability is conducted before, the information obtained proves to be helpful to meet the desired level of testability coverage, but it has less effect on making the design more testable in this paper, different approaches of model testability are described the basic problem with all these approaches are behavioral architecture in model testing. The considered criteria in research can be used as a feature in new systems or it can also be used in helping the selecting model for testability approaches. One single approach cannot be selected in all cases, and decision of a particular approach purely depends on the particular scenario. However, a comprehensive single approach yet has to be found out for all the cases.

Felderer and Fourneret [16] presented the new modern era of technology, where all the data are present on the Internet, so the security of this cloud computing is very important. The biggest challenge we are facing right now is the security of our software systems as we know the changes in system or environment can cause new threats. For this purpose, we have extracted some security regression testing approaches and classified them according to some security regression approach criteria that are abstraction level, security issue, regression testing technique, and tool support and according to evaluation criteria that are evaluated system, maturity of the system, and evaluation measures. After a thorough research, we found out that regarding abstraction level, model-based approach for regression testing of security mechanism and requirement are famous. With reference to regression testing techniques and security issues, we observed that plain re-test all approaches are dominating. Regarding tool support, we found out that these approaches have small tool support for regression testing techniques and more support for automated test execution. While evaluation system techniques for which regression security testing is applied are very rare.

Baride and Dutta [17] introduced a cloud-based model for mobile environment, where actual devices and cloud-based platform have used for system configuration. This model also provides software testing for automatically execution of various tests based on given applications. Cloud computing is recent evaluation for distributed computing, where central remote servers maintain the data and applications. In this proposed model, the various services and different set of testing services of mobile applications have offered. The proposed model has emulator, testing mobile environment and actual devices, automate testing, testing application complexity, and testing for different platforms. In emulator, the duplicate functions for different mobile computing devices are available to speed up the test processes. In testing phase, the applications have tested for all types of mobile devices. In automate testing, the cloud services integrated with many automation tools. This model also offered the testing, security, and synchronization testing facilities by cloud services.

Murugesan and Balasubramanian [18, 19] proposed a cloud testing, where the devices easily allocate the required resources using virtualization and easy to scale up facilities anytime without effecting the entire systems. One of the main advantages of these services is cost effectiveness. There are many challenges for testing the mobile applications such as restricted computational resources and complex testing processes. Device diversity is another challenge where each mobile phone differed in screen orientation, size, and architecture. The author addresses these limitations in this model and offered testing as a service (TaaS). TaaS cloud infrastructure has offered new service model, where devices are integrated with cloud bases services and fulfill the requirements. This model provides TaaS services, where real devices use testing facilities through cloud services like Dalvik cache, iOS, and Android. In addition, the various automated testing tools like Tets, IBMs, RQM, and HP-Quick have offered (Table 1).

Table 1 Software testability its benefits and limitations

Comparison between software testing techniques			
Author	Method used	Efficiency	Description
Wohlin [1]	Snowballing	High	Extraction of knowledge from huge data. Used in repository, databases, and World Wide Web
Garousi et al. [2]	R&D	High	Reduced the failures occur due to disagreement in confidentiality algorithm
Shaheen and Du Bousquet [3]	Metrics	High	Provides easy to use metrics for testability of source code
Ertam and Aydın [4]	Conventional neural networks	Very high	Machine learning for learning purposes which gives good results especially to find the unknown patterns
Huda et al. [5]	Comparison	Medium	Proved that software testability should be done earlier to improve quality and testing time
Häser et al. [6]	Model-based integration technique	Medium	Reduced the upcoming security challenges and issues
Garousi [7]	Relating software performance with software testability		Measured system performance within constraint (speed, accuracy, memory)
Hassan et al. [8]	Checked software testability with different angles	Medium	Found percentage of constraints of software testability and software performance
Felderer and Fourneret [16]	Model-based security testing	Low	Improve the security of particular software
Petersen et al. [12]	Systematic mapping	Medium	Gave list of activities for systematic mapping and made amendments in systematic mapping guideline
Suri and Singhani [13]	Object-oriented testing	Medium	Discussed techniques of implementation of object-oriented testing
Nikfard et al. [15]	Comparison of old techniques	Low	Proved that one single approach cannot be selected for testing of design

(continued)

Table 1 (continued)

Comparison between software testing techniques			
Author	Method used	Efficiency	Description
Felderer and Fourneret [16]	Security regression testing technique	Medium (rare)	Proved that regarding abstraction level model-based approach, regarding security issue plain re-test technique, and regarding tool support automated test execution techniques are famous

4 Conclusion

After a detail review in the field of software testing, we conducted a short and concise. This review paper aims to providing the most comprehensive review in the field of software testability. We reviewed the most recent and selected papers and investigated the papers and reviewed that how the previous studies tackled software testing. Most of the studies are based on different types of methods or techniques related to testability measurement. The most often mentioned factors are controllability and observability. We observed that the most of the previous literatures are on testability measures and discussed how to improve the testability. In addition, literature also discussed the observability and controllability. In future, we will review more literature in the field of software testing and develop a testability framework for testing the non-functional properties such as security and performance.

References

1. Wohlin C (2014) Guidelines for snowballing in systematic literature studies and a replication in software engineering. In Proceedings of the 18th international conference on evaluation and assessment in software engineering, p 38
2. Garousi V, Eskandar MM, Herkiloğlu K (2017) Industry–academia collaborations in software testing: experience and success stories from Canada and Turkey. Software Qual J 25:1091–1143
3. Shaheen MR, Du Bousquet L (2010) Survey of source code metrics for evaluating testability of object oriented systems
4. Ertam F, Aydın G (2017) Data classification with deep learning using Tensorflow. In 2017 international conference on computer science and engineering (UBMK), pp 755–758
5. Huda M, Arya Y, Khan M (2014) Measuring testability of object oriented design: a systematic review. Int J Sci Eng Technol (IJSET) 3:1313–1319
6. Häser F, Felderer M, Breu R (2014) Software paradigms, assessment types and non-functional requirements in model-based integration testing: a systematic literature review. In Proceedings of the 18th international conference on evaluation and assessment in software engineering, p 29
7. Garousi V (2015) A bibliometric analysis of the Turkish software engineering research community. Scientometrics 105:23–49

8. Hassan MM, Afzal W, Lindström B, Shah SMA, Andler SF, Blom M (2016) Testability and software performance: a systematic mapping study. In Proceedings of the 31st annual ACM symposium on applied computing, pp 1566–1569
9. Suri PR, Singhani H (2015) Object oriented software testability (OOSTe) metrics analysis. Int J Comp Appl Technol Res 4:359–367
10. Felderer M, Zech P, Breu R, Büchler M, Pretschner A (2016) Model-based security testing: a taxonomy and systematic classification. Softw Test Verification Reliab 26:119–148
11. Petersen K, Vakkalanka S, Kuzniarz L (2015) Guidelines for conducting systematic mapping studies in software engineering: an update. Inf Softw Technol 64:1–18
12. Petersen K, Feldt R, Mujtaba S, Mattsson M (2008) Systematic mapping studies in software engineering. Ease 68–77
13. Suri PR, Singhani H (2015) Object oriented software testability survey at designing and implementation phase. Int J Sci Res 4:3047–3053
14. Binder RV (1994) Design for testability in object-oriented systems. Commun ACM 37:87–102
15. Nikfard P, Ibrahim SB, Rouhani B, Selamat HB, Mahrin MN (2013) An evaluation for model testability approaches. Int J Comp Technol 9
16. Felderer M, Fourneret E (2015) A systematic classification of security regression testing approaches. Int J Softw Tools Technol Transfer 17:305–319
17. Baride S, Dutta K (2011) A cloud based software testing paradigm for mobile applications. ACM SIGSOFT Softw Eng Notes 36:1–4
18. Murugesan L, Balasubramanian P (2014) Cloud based mobile application testing. In 2014 IEEE/ACIS 13th international conference on computer and information science (ICIS), pp 287–289
19. Obaid AJ, Chatterjee T, Bhattacharya A (2020) Semantic web and web page clustering algorithms: a landscape view. EAI Endorsed Trans Energy Web 8(33)

Development of Sign Language Recognition Application Using Deep Learning

N. R. Rajalakshmi

Abstract Deaf and dumb people use a sign language that can only be communicated through hand gestures to express their ideas and views. This coded language is mainly used by people who have speech and/or hearing impairment. The sign language is constructed by various movement of hands, arms, legs, or facial expressions to express their opinions. Meanings are communicated for every movement or position of gesture. Hand gesture plays a significant role to make mother tongue of impairment people for daily communication. The captured image feature can be extracted to translate the hand gesture communication to text\voice format to minimize the gap between the deaf and normal persons. This work considers the images of sign numerals to classify the numbers 0–9 and the alphabets for A–Z (including space).

Keywords Gestures · Camera sensors · Convolutional neural network · Deep learning

1 Introduction

IoT is used for integration of multiple devices which communicate, interact, and sense with their internal and external states through the embedded technology. IoT offers suitable solutions for several applications such as smart cities, security, traffic congestion, industrial control, and agriculture. The growth of these technologies in the twenty-first century helps to the development of sign language to meet the need of impairment people whose needs are essential. So that, the gap between the specially disabled section and privileged section of the society can be diminished to some extent. The IoT forecasts the route for creation of sign language system to help in a great extent to the people who are having speech and/or hearing impairment can easily express their feelings and emotions. The IoT technology becomes universal, because IoT devices are designed with low-cost component like microprocessors, memory,

N. R. Rajalakshmi (✉)
Computer Science & Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai 600062, India
e-mail: drnrrajalakshmi@veltech.edu.in

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_24

sensors of light, thermal, magnetometer, gyroscope, microphone, accelerometer to grab the user's emotional and behavioral information as well as contextual information. This information is transferred to the cloud platform for analysis via the wireless communication technologies.

Hence, the IoT technologies are exploited here to sense the gesture position. The sensed data which are communicated for processing and analyzed with knowledge base to develop a code language. The gesture recognition mostly uses the digital cameras, gyroscopes, or accelerometers sensors which often integrated inside the mobile or wearable devices. The wearable devices can be programmed to translate the sign language to voice format. The models that were designed use the concepts of Internet of things (IoT) or computer vision for its accomplishment. A keen observation shows that the former process is costly and requires adequate maintenance to establish its proper functioning while the later was effective to either identify the digits or the alphabets individually. The main idea of this work is to develop an effective and robust hand gesture recognition model that will be able to identify the gestures from 0 to 9 and A–Z (including space). The concepts of deep learning used to classify the captured hand sign to provide the output as the respective alphabet or number.

1.1 Literature Survey

Many authors discussed about building a robust real-time hand gesture detection model to identify the various vision-based hand gestures with the help of visual information from the camera [12]. The gestures on a complex back ground like variation in lighting and interposition of obstacles can be detected accurately using depth information from binocular vision system [3]. This system builds the three-dimensional map to detect the gesture position accurately. The researcher also developed the recognition system using magnetometer and infrared sensor technology. An accelerometer-based pen device is also presented in [13] to design a user independent hand gesture recognition system. The researcher also proposed the ambient sensors or wearable sensors for gesture recognition. The sensory data that are perceived from the background sensors used for analysis are very effective one in control the appliance at home [14]. While the later uses the smart watch, smart gloves, or mobile data by considering four most critical gesture analysis of posture, position orientation, and motion. In [4], a pair of gloves using wearable technology are also presented to convert hand gestures to text and speech, respectively. The person who is having disability can wear this glove to have a daily basis communication. The Raspberry Pi3 as the microcontroller which is coupled with flex sensors and accelerometer sensors to recognize the hand gesture. In [14], a wrist worn device is discussed about the capture of 3D pose of the hand and free hand interaction to recognize the gesture. Since the growth of mobile technologies are extremely high, the sensors embedded in the mobile devices are also exploited in various study to continuously capture the data streams of user's gestures. The reinforcement learning-based model also proposed to classify the surgical gesture, wherein the system performs the action based on learning from analysis of video images in the environment. But live video streaming process is challenging one to identify the start and end of any particular video. The authors of [5] discussed about building a robust real-time hand gesture detection model that is capable of detection of hand, identifying various gestures with the help of skin segmentation and hand tracking procedures by using the bootstrap classifiers.

In this work, the various patterns of the user's gesture images are collected. Then, the gesture positions are preprocessed, segmented, and features are extracted and classified.

1.2 Proposed System Architecture

More reliable classification system was developed that will be able to identify the gestures from 0 to 9 and A–Z. The discussed process would reduce the confusion among similar digits and alphabets as well as this process requires zero maintenance. The architecture shown in Fig. 1 depicts the overall working of the system.

The process starts with collecting the gesture position. After which, it is preprocessed and extracted the required features from the gesture position by using convolution neural network. The gesture position is collected either by using Web camera or datasets of 0–9 and A–Z (Source: Kaggle) is used. A sample image of the class 0 is shown below in Fig. 2.

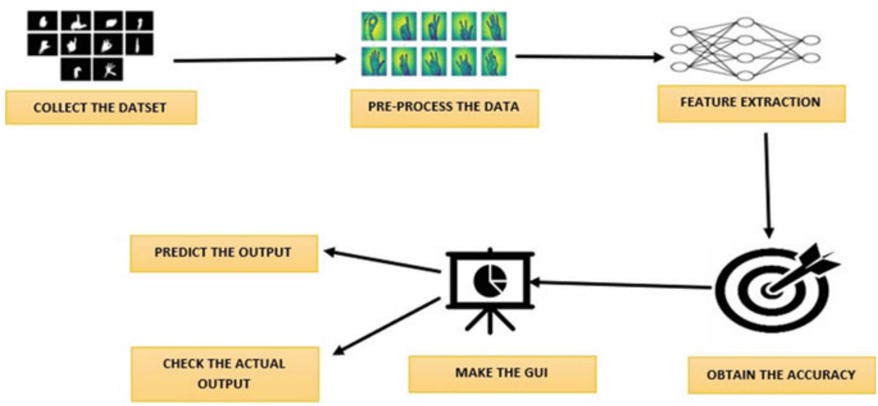


Fig. 1 Sign language recognition model

Fig. 2 Sample dataset

2 Data Preprocessing

The main aim of this module is to preprocess the dataset, so that the data can be effectively transformed and manipulated as per the requirements and thereby yield the most efficient result when any algorithms are applied to it. Here, the concept of convolution neural network is used to extract the feature of gesture position (Fig. 3).

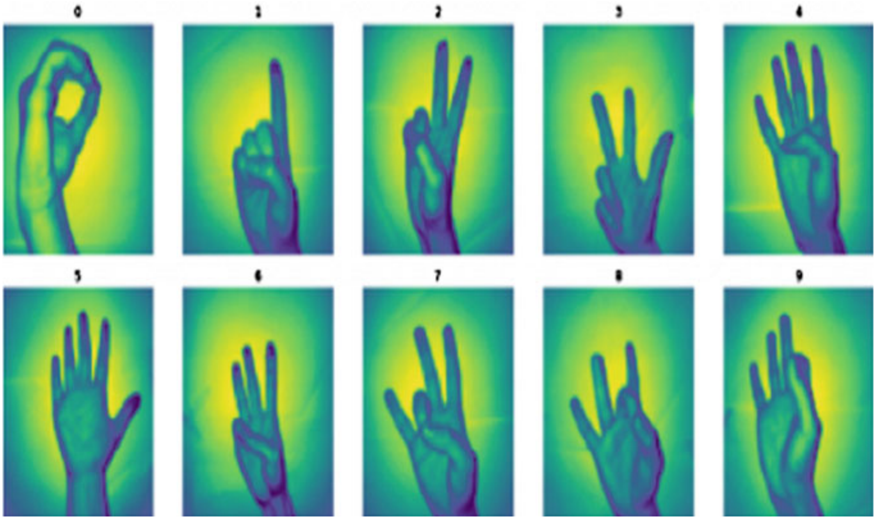


Fig. 3 Preprocessing the data

2.1 Feature Extraction

Feature extraction can be defined as a type of dimensionality reduction where a large quantity of pixels of images is efficiently represented in order to focus more on the interesting portion of the image as compared to the whole image. These help in obtaining a more accurate result. Artificial neural networks (ANNs) with multiple layers are referred to as deep learning or deep neural networks. It has been regarded as one of the most powerful tools in recent decades, and it has been quite popular in the literature due to its ability to handle large amounts of data. The desire for deeper concealed layers has recently begun to outperform traditional solutions in certain situations. CNN is a widely used supervised learning model with a wide range of adaptability performs admirably and particularly impressive. It is frequently used to handle picture data. The outline of all the elements and significant topics associated to CNN is described here, as well as how these elements interact.

2.2 Basic CNN Components

The core components of CNN are extremely similar. The example of LeNet-5 is shown in Fig. 4. Network is made up of three types of layers: convolutional, pooling, and fully-connected. The convolutional layer's goal is to learn the input feature representations. The convolution layer is made up of many convolution kernels that are utilized to generate various feature maps. All neurons in the previous layer are connected to every single neuron in the current layer to extract the semantic information. The convolving the input with a learnt kernel may create new feature and then applying an element-wise nonlinear activation function on the convolved outputs. ReLU is a regularly used activation strategy in CNN to prevent vanishing gradients and speedup the training.

Several different kernels are used to create the entire feature maps. The first convolutional layer's kernels are trained to mine the low-level features like edges and curves, while higher-layer kernels are trained to encode more abstract features

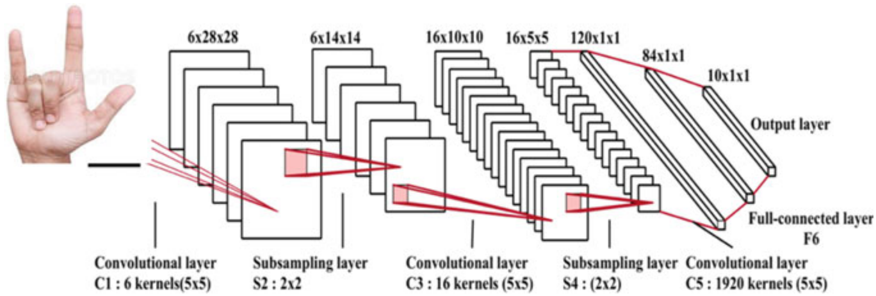


Fig. 4 LeNet-5 network

of region of interest of gesture is depicted. Several convolutional and pooling layers are stacked to achieve high-level feature representation. CNN emphasizes the concept of pooling. It reduces the number of connections between convolutional layers, which reduces the computational cost. Pooling can be done in three different ways: minimum, maximum, and average. The pooling layer is one of the most important network layers since it prevents overfitting and allows the network to be used as a basic model. After numerous convolutional and pooling layers that try to execute high-level reasoning [9, 11, 17] one or more fully-connected layers may be there. It is worth noting that a 1×1 convolution layer can be used instead of a fully-connected layer at the output layer. The active function of softmax is simply a probability function that is used when there are a lot of classes in a multiclass function.

3 Implementation

Implementation is a vital step for understanding the proposed model effectiveness. It is more like incorporating and carrying out the steps in a sequential in order to fetch the desired output (Fig. 5).

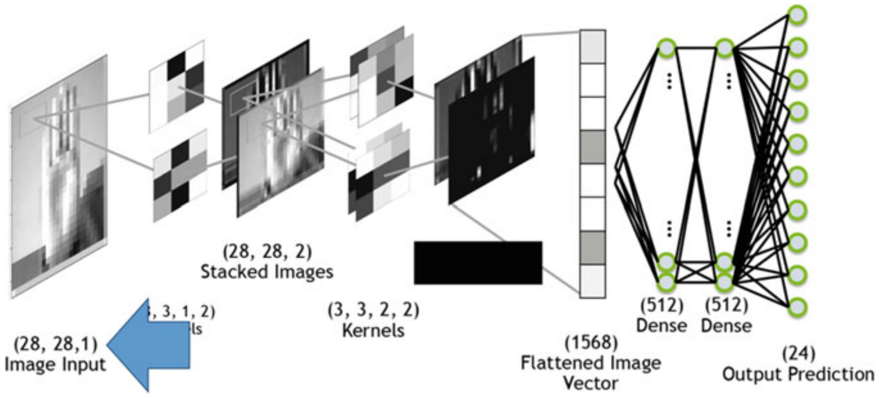


Fig. 5 Convolutional model

The overall model of the proposed CNN is given above. The dataset is divided as training and validation dataset. The reshape of the dataset is done, so that they are in a 28×28 pixel format. This will allow the convolutions to associate groups of pixels and detect important features. The above diagram shows the layers of 2D convolutional. Small kernels will comb over the input image and find more complex features that are relevant to classification. Let us look at our first Conv2D layer:

model.add(Conv2D(75, (3, 3), strides = 1, padding = 'same' …)

Table 1 CNN performance analysis using Adam optimizer

No. of runs (epoch)	Training accuracy	Validation accuracy	Training loss	Validation loss
10	99.9	92.07	0.003	0.55
15	99.96	96.26	0.0029	0.038
20	99.97	94.2	0.0012	0.32
25	99.97	96.32	0.0012	0.31

75 refers to the number of filters that will be learnt. (3,3) refers to the size of those filters. Strides are the size of the steps taken by the filter as it passes over the image. Padding refers to whether the output image from the filter will be the same size as the input image. Batch normalization which scales the values in the hidden layers to improve training. Max pooling essentially shrinks the image to a lower resolution. It does this to make the model to be robust to translation and also makes the model faster. Overfitting can be avoided by using a technique called dropout. Dropout picks a subset of neurons at random and shuts them off. So that they do not participate in forward or backward propagation in that pass. This ensures that the network is strong and redundant. Flatten takes the output of one layer which is multidimensional and flattens it into a one-dimensional array. The output is called a feature vector and will be connected to the final classification layer. The first dense layer (512 units) takes the feature vector as input and learns which features will contribute to a particular classification. The second dense layer (24 units) is the final classification layer that outputs the prediction. It looks like this model is significantly improved. The training accuracy is very high, and the validation accuracy has improved as well. The proposed state-of-the-art CNN was constructed and evaluated using the parameters stated below in Table 1 (Fig. 6).

An epoch is basically a hyperparameter that defines the number of times the learning algorithm will work through the entire dataset. The performance metrics of the model using CNN are obtained by running the model till epoch 25.

The training and validation accuracy for the model are also given in Table 1. From the data, it can be deduced that increasing the epoch leads to an increase in accuracy as well as a decrease in loss when compared to the previous iterations. Table 1 displays the effects of running the model for epochs from 1 to 25, and the hyperparameter is adjusted to improve the efficiency with the Adam optimizer.

4 Results and Discussion

It is observed that people are more driven toward the power of technological advancements for performing their day-to-day activities. This model is developed keeping the fact in mind that all the people across the Globe irrespective of their caste, creed, and disabilities should enjoy the technological advancements, and it should not be

```
[8]: model.summary()

Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 28, 28, 75)        750
_________________________________________________________________
batch_normalization (BatchNo (None, 28, 28, 75)        300
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 14, 14, 75)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 14, 14, 50)        33800
_________________________________________________________________
dropout (Dropout)            (None, 14, 14, 50)        0
_________________________________________________________________
batch_normalization_1 (Batch (None, 14, 14, 50)        200
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 7, 7, 50)          0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 7, 7, 25)          11275
_________________________________________________________________
batch_normalization_2 (Batch (None, 7, 7, 25)          100
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 4, 4, 25)          0
_________________________________________________________________
flatten (Flatten)            (None, 400)               0
_________________________________________________________________
dense (Dense)                (None, 512)               205312
_________________________________________________________________
dropout_1 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_1 (Dense)              (None, 24)                12312
=================================================================
Total params: 264,049
Trainable params: 263,749
Non-trainable params: 300
```

Fig. 6 Summary of the model

confined to any particular section of the society. The gesture of number 4 is shown in Fig. 7.

Here, the various hand gestures ranging from 0 to 9 and A–Z (including space) were thoroughly studied made model for creating user friendly with the people. People require zero knowledge and no experience along with little maintenance to use it. Additionally, since every gesture used is unique, words, sentences, and even paragraphs can be formed using this model, thereby enabling the specially challenged people to convey their thoughts and feelings with other people. Figure 8 represents the sample output obtained for class 'C'.

5 Conclusion

The gesture recognition classification platform for classifying the hand gestures between 0–9 and A–Z (including space) is successfully built. All the gestures across all the classes were unique, and there were no confusions in between any two gestures.

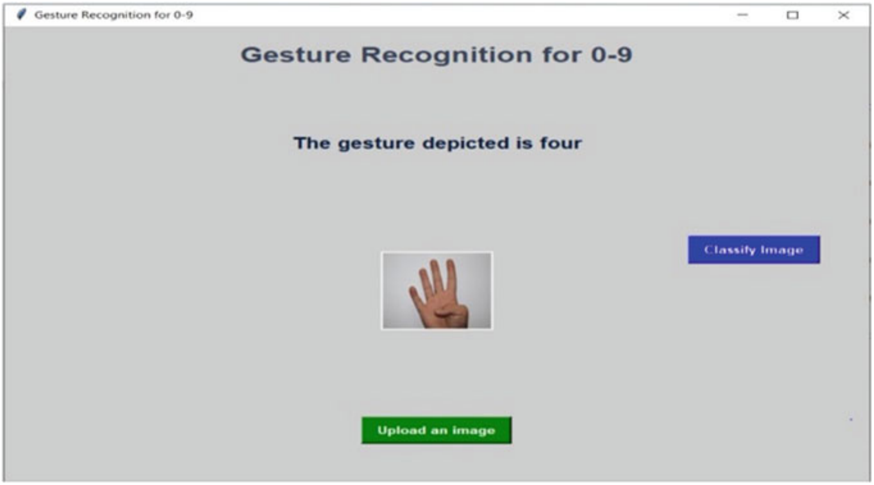


Fig. 7 Depiction of gesture for number 4

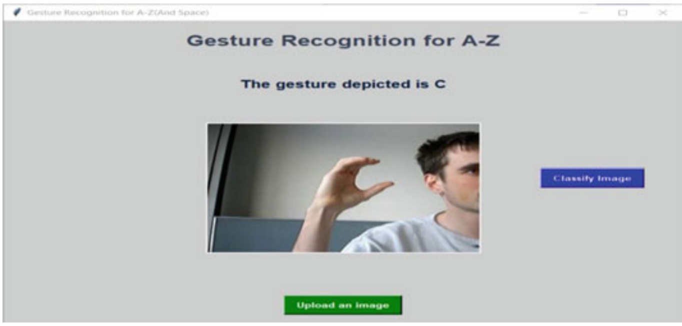


Fig. 8 Depiction of gesture for alphabet C

An accuracy of 99.97% was obtained. Although the model has been built, acquired a high percentage of accuracy and is able to do its job of classifying the images effectively, but the process may be time consuming when paragraphs containing hundreds of words have to be considered. Therefore, the proposed system can still be modified in future in order to build a more robust system.

References

1. Jin CM, Omar Z, Jaward MH (2016) A mobile application of american sign language translation via image processing algorithms. In: IEEE Region 10 symposium (TENSYMP), Bali, Indonesia
2. Gao X, Jin Y, Dou Q, Heng PA (2020) Automatic gesture recognition in robot-assisted surgery

with reinforcement learning and tree search. arXiv preprint, arXiv:2002.08718
3. Jiang D, Zheng Z, Li G, Sun Y, Kong J, Jiang G, Liu H (2019) Gesture recognition based on binocular vision. Clust Comput 22(6):13261–13271
4. Vijayakumar KP, Nair A, Tomar N (2020) Hand gesture to speech and text conversion device. Int J Innov Technol Explor Eng (IJITEE) 9(5)
5. Kopuklu O, Gunduz A, Kose N, Rigoll G (2019) Real-time hand gesture detection and classification using convolutional neural networks. arXiv:1901.10323v3 [cs.CV]. Accessed 18 Oct 2019
6. Cheok MJ, Omar Z, Jaward MH (2017) A review of hand gesture and sign language recognition techniques. Springer, GmbH Germany
7. Liang RH, Ouhyoung M (1998) A real-time continuous gesture recognition system for sign language. In: IEEE international conference on automatic face and gesture recognition, pp 558–567, Japan
8. Gurav RM, Kadbe PK (2015) Real time finger tracking and contour detection for gesture recognition using OpenCV. In: 2015 international conference on industrial instrumentation and control (ICIC), College of Engineering Pune, India, May 28–30
9. Ravikiran J, Mahesh K, Mahishi S, Dheeraj R, Sudheender S, Pujari NV (2009) Finger detection for sign language recognition. In: Proceedings of the international multi conference of engineers and computer scientists 2009 Vol I IMECS 2009, March 18–20, 2009, Hong Kong
10. Rumyantsev O, Merati M, Ramachandran V (2012) Hand sign recognition through palm gesture and movement image processing. EE 368, Spring 2012
11. Rekha J, Bhattacharaya J, Majumder S (2011) Hand gesture recognition for sign language: a new hybrid approach. In: International conference on vision and pattern recognition, IPCV'11
12. Singha J, Roy A, Laskar RH (2018) Dynamic hand gesture recognition using vision-based approach for human–computer interaction. Neural Comput Appl 29(4):1129–1141
13. Xie R, Cao J (2016) Accelerometer-based hand gesture recognition by neural network and similarity matching. IEEE Sens J 16(11):4537–4545
14. Venkatnarayan RH, Shahzad M (2018) Gesturerecognition using ambient light. Proc ACM Interact Mob Wearable Ubiquitous Technol 2(1):40, 1–40

Changing Many Design Parameters in the Performance of Single-Sided Linear Induction Motor (SLIM) for Improved Efficiency and Power Factor

Hayder Hussein Kadhum, Hayder H. Enaw, and Karrar M. Al-Anbary

Abstract Linear induction motors (LIM) are applied obviously in industrial applications, mainly in linear motion models. In fact, these machines have two general problems which are low efficiency and low power factor. These issues cause different side effects such as high power consumption, high input current, and occupation. This research studies the dynamic behavior of a single-sided linear induction motor by changing the basic design parameters. The process of selecting the equivalent circuit elements was also carried out using a new algorithm to obtain a modified model to meet the requirement, which is to improve the efficiency and power factor simultaneously. The final equivalent circuit of the modified model was adopted as an input to a simulation program using MATLAB/SIMULINK, which helped build an integrated model using mathematical equations with the specified inputs and the required outputs. Also, the process of testing and comparing the results is implemented for the proposed model in different cases with and without load. Then, analysing the results of motor performance and response speed using a voltage source inverter (VSI) with and without using a PID controller to improve the dynamic performance of the model, and speed control. Finally, the optimization results are validated, and the results are compared. The outcomes of proposed algorithm are shown by the genetic algorithm (GA), particle swarm optimization (PSO), and Cuckoo search with respect to the power factor and efficiency.

Keywords Efficiency and power factor improvement · Analytical model · (SLIM) design · Thrust

H. Hussein Kadhum (✉) · H. H. Enaw · K. M. Al-Anbary
College of Engineering, Electrical Engineering Department, University of Babylon, Babel, Iraq
e-mail: eng.hayder.kadhum@uobabylon.edu.iq

H. H. Enaw
e-mail: eng.hayder.hamzah@uobabylon.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_26

1 Introduction

During last decades, electrical systems started using SLIM obviously in transferring electrical power. This is because SLIM has important features as compared with rotary induction motor (RIM) which are higher generating power without mechanical contacts, longer life of wheels, better road line, lower radius of turns, and quicker deceleration as well as acceleration [1]. However, in many applications, which need linear motion, linear induction motors (LIMs) represent the best option due to their properties such as cheap price and maintenance, high deceleration or acceleration, and easy design. As a result, LIMs have been applied largely in different applications like transferring systems, processing materials, smart applications such as opening and closing doors and others automatically and so on [2, 3].

Charles Wheatstone designed first linear motor in 1845 [5–7]. During last decades, SLIM has been studied obviously in different analysis techniques. However, there are some reasons which make analysis of this motors type very difficult. The first main reason is the disability of identifying size of the electrical and magnetic loads. This prevents determining apparent power. One of the main steps in analyzing with single-sided LIM with a solid-steel reaction plate is calculating its mutual reactance. It is determined by two-dimensional field distributions between the air-gap as well as the secondary windings. The 2D FEM of field distribution is applied to find the parameters of a two-axis design. The static nonlinear vector potential technique has been used in determining flux distribution with iron at saturation. However, the linear time harmonic vector potential field is applied when computing inductance with analytical model. These inductances help effectively in determining the power factor in addition to the efficiency [8]. For developing LIMs response, the control as well as the design of these motors should be optimized continuously. Most of the researches ignore optimizing design for LIMs [9]. Generally, the researches focus on expenditure, starting torque, primary weight, and winding design, while other important factors, like efficiency as well as power factor (PF), did not take enough concentration. Furthermore, low efficiency means higher energy consumption. Additionally, low PF leads to more power losses during transferring. As a result, inverter cannot work optimally. In general, there are several main features which are adopted in optimization design like transverse edge influence, end effect as well as normal force aspects [10–13].

Moreover, end effect affects directly on the efficiency as well as power factor and that leads to change design features of SLIM as well. Furthermore, the structure of the motor changes its pole pitch. Also, identifying frequency and rated slip depends mainly on the rated speed of the motor [14–18]. This paper gives a modern technique depending on the air-gap flux density equation. This equation does not ignore the influence of end effects. As a result, several nonlinear equations have been built which satisfy highest efficiency with smallest primary weight. After that a suitable method is required for calculating independent variables like the ratio between slot width and slot pitch, number of poles, stack height, pole pitch, and so on [14–16].

The goal of this research is to study the performance of a modified model of SLIM by varying its parameters to reduce the power consumptions and increase the power factor. The design parameters and the equivalent circuit are used as optimization technique. They help effectively in enhancing efficiency as well as power factor of the final proposed model. This model is tested and simulated thorough applying MATLAB/SIMULINK programming. The characteristics and performance of this model are studied for different load conditions. For instance, it is connected directly to the supply voltage and then through voltage source inverter (VSI) with and without the aid of PID controller.

.

2 Physical Structure and Magnetic Equivalent Circuit Models

Figure 1a and b below demonstrates the structure of the short stator and rotor of LIM which is used in this study.

In general, the changing of skin at rated frequency is very little. This is because the conductive sheets on secondary side are very slim. As a result, equivalent secondary inductance in the proposed design is ignored the resistance of any phase of the LIM stator windings. R_1 is measured by the following formula:

$$R_1 = \rho_w \frac{l_w}{A_w} \tag{1}$$

while ρ_w is the wire resistivity of stator winding, l_w is wire length for each phase separately, and A_w is cross-section area of the wire. This leakage reactance can be calculated as:

$$X_1 = \frac{2\mu_0 \pi f\left[\left(\lambda_s\left(1+\frac{3}{p}\right)+\lambda_d\right)\frac{w_s}{q_1}+\lambda_e l_{ce}\right]N_1^2}{p} \tag{2}$$

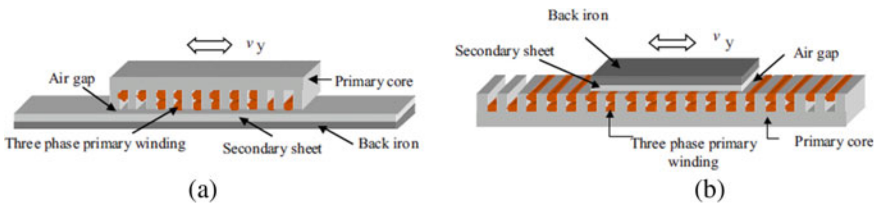


Fig. 1 Structures of LIM **a** short stator **b** short rotor

$$\lambda_s = \frac{h_s(1+3k_p)}{12w_s}, \ \lambda_e = 0.3(3k_p - 1) \text{ and } \lambda_d = \frac{5\left(\frac{g_e}{w_s}\right)}{5+4\left(\frac{g_0}{w_s}\right)}$$

The magnetizing reactance X_m is shown in

$$X_m = \frac{24\mu_0 \pi f W_{se} k_w N_1^2 \tau}{\pi^2 p g_e} \tag{3}$$

where W_{se}, stator width given as

$$W_{se} = W_s + g_o \tag{4}$$

The goodness factor is defined as:

$$G = \frac{2\mu_0 f \tau^2}{\pi\left(\frac{\rho_r}{d}\right) g_e} \tag{5}$$

3 Thrust and Efficiency

Generally, electromagnetic torque as well as rotation speed represents the main output parameters of the rotary motors which are measured by N, M, and Rad/sec, respectively, while electromagnetic thrust and linear speed are mostly used as output elements of linear motors that measured by N and m/sec sequentially. Furthermore, the following formula is applied to calculate electromagnetic thrust [18].

$$P_o = mI_2^2 \frac{R_2}{S} - mI_2^2 R_2 = mI_2^2 R_2 \left(\frac{1-S}{S}\right) \tag{6}$$

$$P_o = F_s v_r \tag{7}$$

$$v_r = (1-S)v_s \tag{8}$$

$$F_s = \frac{mI_2^2 R_2}{v_s S} \tag{9}$$

By putting Eq. (8) in Eq. (9) the SLIM electromagnetic thrust becomes:

$$F_s = \frac{mI_1^2 R_2}{\left[\frac{1}{(SG)^2}\right] v_s S} \tag{10}$$

Then, the efficiency as well as power factor of LIM will be as following:

$$\eta = \frac{F_s 2\tau f_1 (1 - S)}{F_s 2\tau f_1 + 3R_1 I_1^2} \tag{11}$$

$$\cos\phi = \frac{F_s 2\tau f_1 + 3R_1 I_1^2}{3V_1 I_1} \tag{12}$$

There are various factors effect on efficiency and power factor. Therefore, applying optimization technique between these factors helps effectively in identifying the suitable parameters.

4 Voltage Source Inverter (VSI) Technique

Inverters are the devices which are changing the electrical power from DC to AC. In these devices; the output power is taken with a specific voltage value as well as frequency. During voltage source inverter (VSI) technique, the output voltage can be varied with a specific frequency and vice versa. Changeable output voltage is determined through changing the input DC voltage or by gate signal of the inverter, which is normally accomplished by pulse width modulation (PWM) technique [21, 22]. The major reason of applying these methodologies is to supply a three-phase voltage source to LIM. In fact, the value and frequency of the voltages must be convenient. This helps greatly in controlling the speed of LIM easily. MATLAB program is used in simulating the performance of SLIM. Stationary reference frame helps largely in modeling the suggested motor. The block diagram of SLIM fed from VSI is shown in Fig. 2. The inverter phase voltages that fed to SLIM with modulation index equal to 1 and 0.5 are shown in Figs. 3 and 4, respectively.

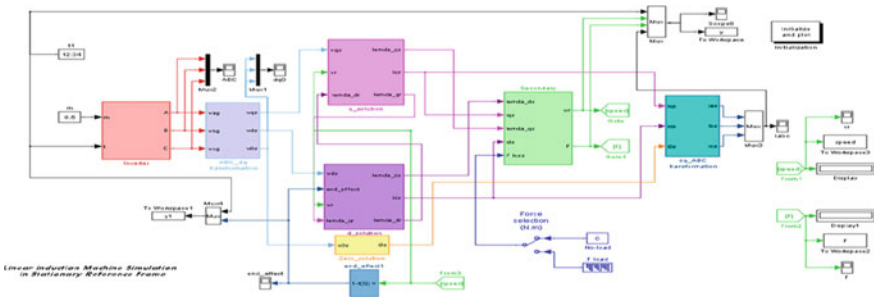


Fig. 2 Linear induction motor simulation using stationary reference frame fed from VSI

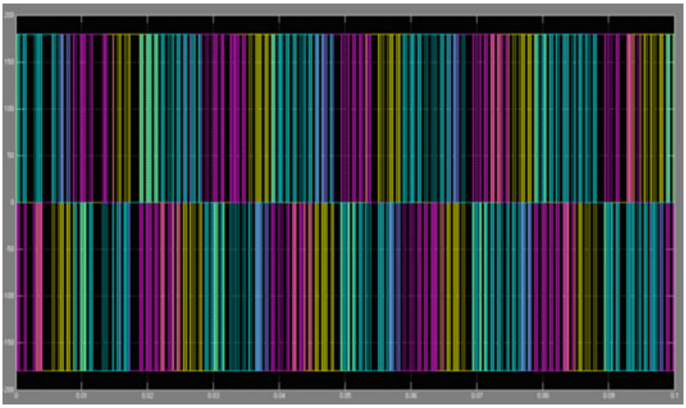


Fig. 3 Inverter voltage at modulation index = 1

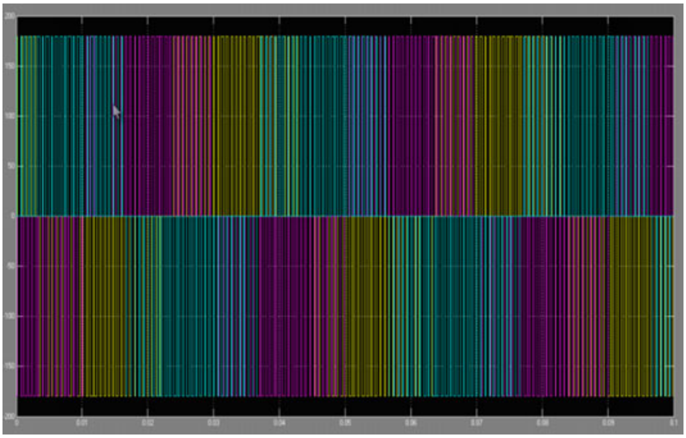


Fig. 4 Inverter voltage at modulation index = 0.5

5 SLIM Design and Optimization

For enhancing the performance of SLIM, the elements which have major influence must be identified. For this model, these elements are efficiency, power factor, maximum thrust slip, aluminum thickness, primary width or pole pitch, primary current density, and the braking force on the end effect. In this paper, several design factors of a reference SLIM design are used to investigate the dynamic performance of this motor. After that the optimization technique is applied to achieve suitable equivalent circuit of the modified suggested system. This methodology is achieved by developing and using a computer program based on MATLABfor optimization and enhancement evaluation, for the efficiency and power factor product. For evaluating the response of SLIM, the suggested elements are changed separately. Then, the optimal value for every factor is recognized. The analysis process is implemented through altering one factor with keeping the others constant. After that the effect of varying these parameters on the performance of SLIM is analyzed, and the results are discussed. The suggested design has to involve the optimal value of used factors for satisfying the higher efficiency and best power factor. The design steps and evaluation process of the SLIM analysis are summarized in the flowchart shown in Fig. 5. A computer program is developed to implement and processing the motor parameters and operation factors.

6 Simulation Results Using PID Controller

PID is a short of proportional–integrative–derivative. It is the best technique of controllers for processing the errors of data. In general, this method depends on the feedback data of the controller in overcoming the mistakes. In fact, a comparison has been performed between input data and the reference data to find the mistakes. After that the error has been adjusted before input to the system. In some situations, the optimization must be implemented for calculating and processing the outcomes to be suitable. The operating formula of PID controller is shown below:

$$U(t) = K_{\mathrm{p}} e(t) + K_{\mathrm{t}} \int e(t)\mathrm{d}t + K_{\mathrm{D}} \frac{\mathrm{d}}{\mathrm{d}t} e(t) \tag{13}$$

While $u(t)$ represents the output comes of PID controller, K_{P}, K_{I}, and K_{D} are the proportional, integrative as well as derivative gains, correspondingly. Finally, $e(t)$ is the error of the wave which is measured by the following equation:

$$e(t) = r(t) - y(t) \tag{14}$$

$r(t)$ represents the reference value, while $y(t)$ is the output of the circuit. The determination of these three parameters can be tuned by "Trial and Error" principle to get the best controller response. After many testing and running in the simulation, the final best parameters values are as follows:

- Proportional gain (k_{p}): 0.2.
- Integral gain (k_{I}): 60.
- Derivative gain (k_{d}): zero (in this project).

Through changing the gains of the PID controller, the performance of this controller arrives its optimal case at specific values [10]. In fact, proportional gain represents the most important factor. It plays an important role in varying its value with respect to input error. The output of the controller changes according to the

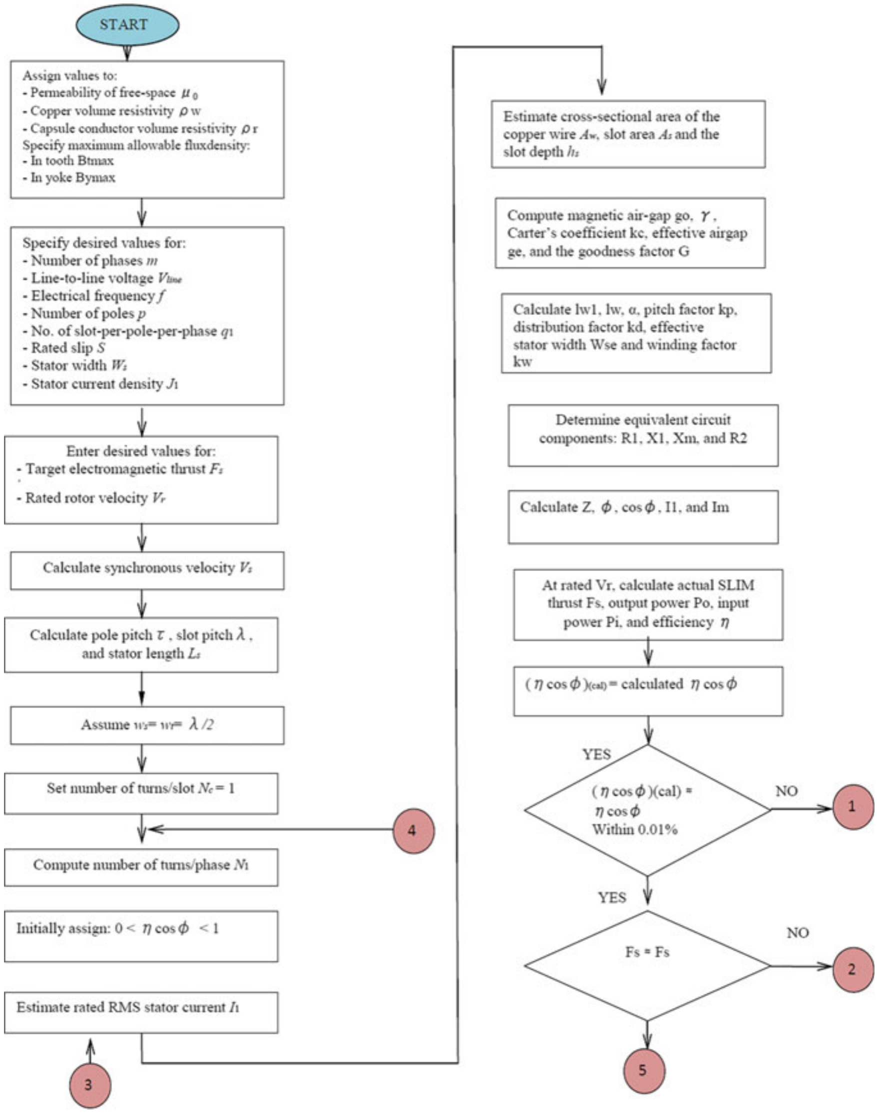


Fig. 5 Flowchart optimizations were implemented depending on the analytical system of the machine, and these data are given in Table 1

proportional gain which depends directly on the error of the input signal. In some time, the value of proportional gain increases largely and then leads to instable situation. Control block implements a PID controller which is reasonable of controlling the reference speed signal in different speed ranges (from zero upward rated speed).

Figure 6 shows the overall diagram of SLIM in MATLAB/SIMULINK using PID controller. In the beginning, KI and KD have been made zero, while KP is varied

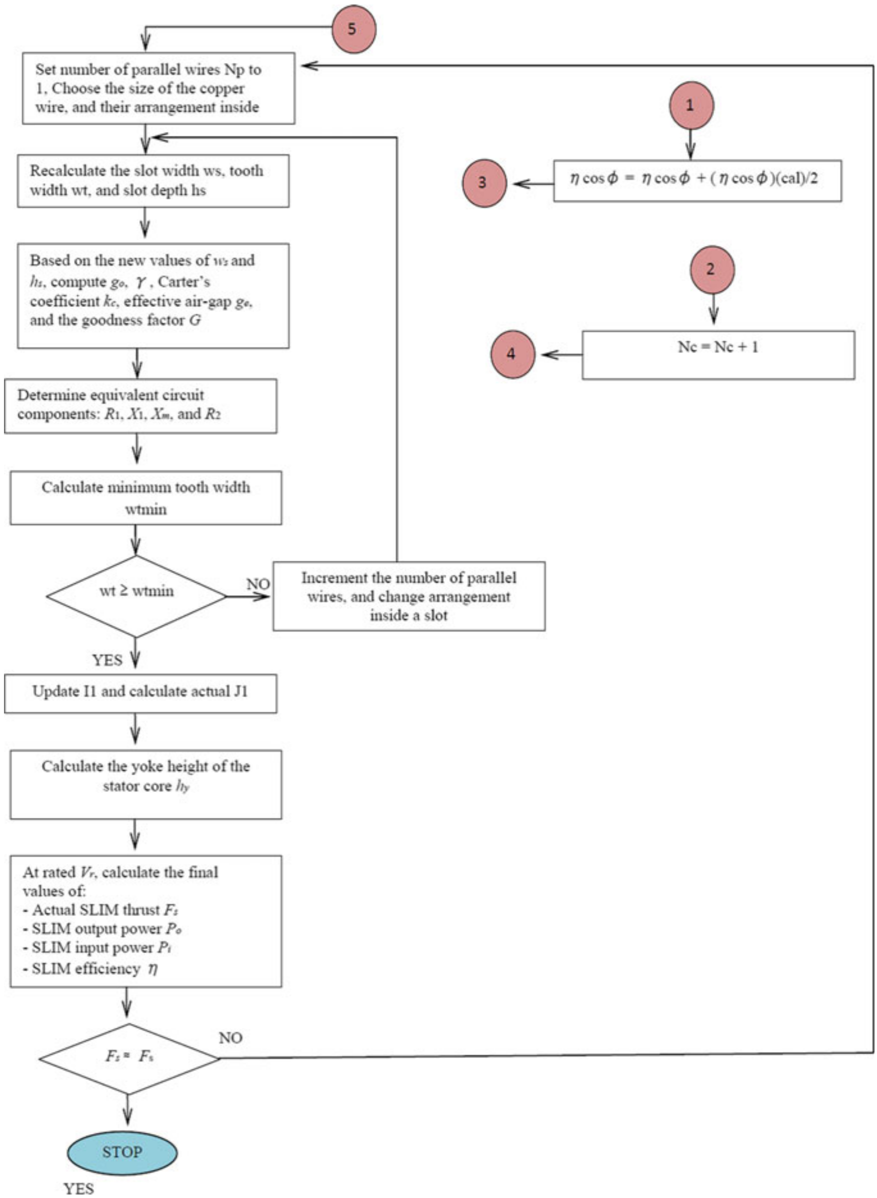


Fig. 5 (continued)

Table 1 Optimized motors specification

Parameter name	Primary length (mm)	Primary depth (mm)	Slot width (mm)	Slip frequency)	Slot depth (mm)	Number of poles	No. of slots	Conductor area (mm^2)	Base velocity (Km)	Efficiency	Power factor
Original	1950	98.6	15.87	6	58	6	80	67.19	40	0.65	0.50
Optimal-1	2341	119.6	19.3	6	83.6	8	78	118.8	40	0.701	0.48
Optimal-2	2333	114.6	19	5.87	79.1	8	78	106.4	40	0.704	0.4657
Optimal-3	2333	114.6	18.8	5.7	74.2	8	78	101.83	40	0.709	0.478

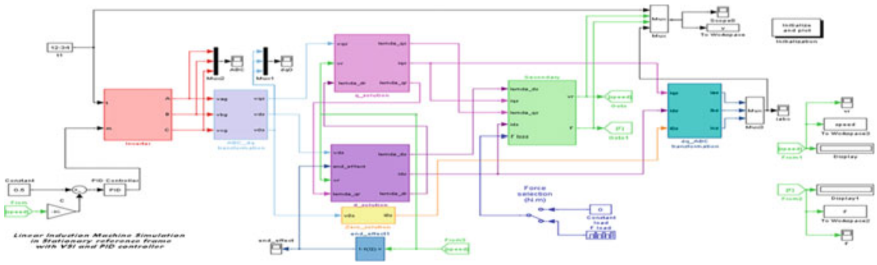


Fig. 6 Block diagram of SLIM with PID controller

in steps, and the response of SLIM is checked in each step until oscillation. Then, KP is made half of oscillation magnitude according to "quarter amplitude decay" type reaction. After that KI will be increased, with saving other gains constants, until overcoming any error in the response. Similarly, increasing KI too much will lead to unstable situation. Lastly, KD will be changed if the response of the system still unacceptable. KD helps the system to satisfy its normal performance in short time after any load disturbance. In the same way, KD can cause overshoot in the system performance if it increases so high.

7 Result

Table 2 below shows the suggested values of best design parameters than other methods (Conventional motor parameters, Cuckoo method, and PSO method) and thus improves the power factor and efficiency.

Table 2 Comparison of various optimization results

Parameter name	Conventional design [17]	Cuckoo search [16]	PSO [15]	Proposed model
Efficiency (%)	36	49.01	54.68	69.94
Maximum thrust slip	0.5	0.5	0.3	0.69
Power factor	0.32	0.5729	0.58	0.711
Primary current density (Amm-2)	4	3	6.5	7.6
Primary width/pole pitch (mm)	2	2	4	5.9
Aluminum thickness (inches)	2	1	4	5.3

Fig. 7 Variation of $\eta \cos \varphi$ product with v_r for different (d)

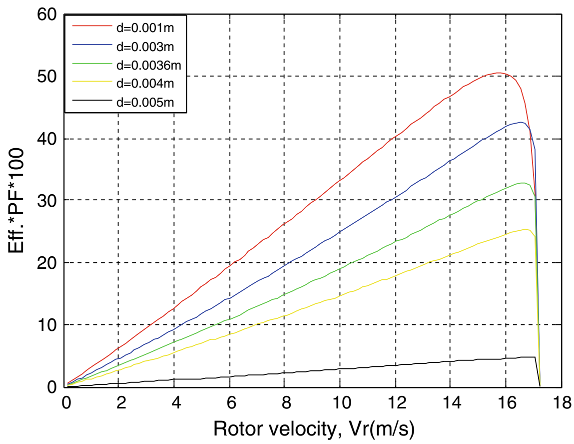


Fig. 8 Variation of $\eta \cos \varphi$ product with v_r for different air-gap length (g_m)

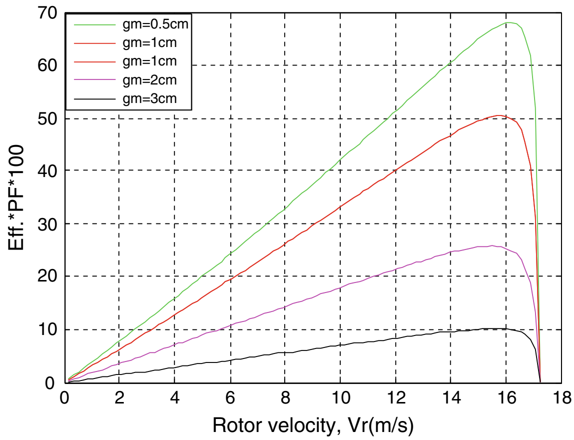


Figure 7 gives the clearly applied result between efficiency with respect to rotor speed of different thicknesses. In addition, this approach is repeated again at different air-gap length, different number of poles, and different slips. Figures 8, 9, and 10 below show these relationships observably.

7.1 Simulation Results for Speed Control Using VSI

Figures 11 and 12 show the thrust force and secondary speed v_r at variable load and at no load, at (modulation index 0.5 and modulation index 1). It is clear when comparing that the speed is reduced to about 7.8 m/s for modulation index = 0.5 as it was 15.7 m/s for m=1. Similar effect of modulation index on v_r is occurring for load

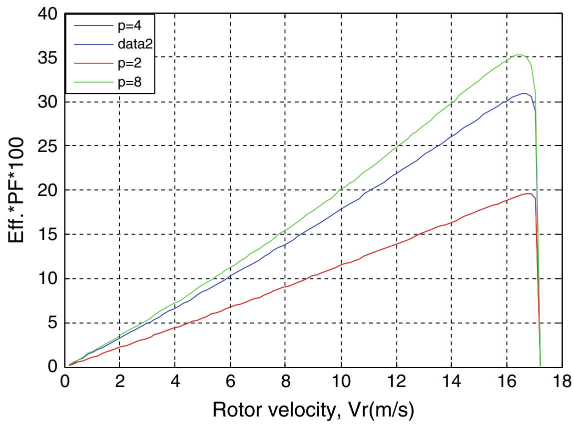


Fig. 9 Variation of $\eta \cos \varphi$ product with v_r

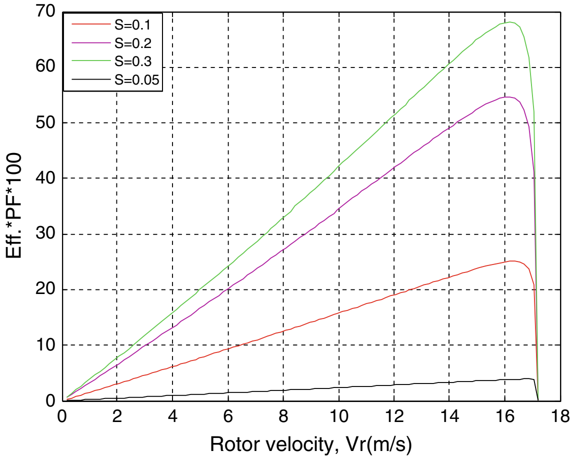
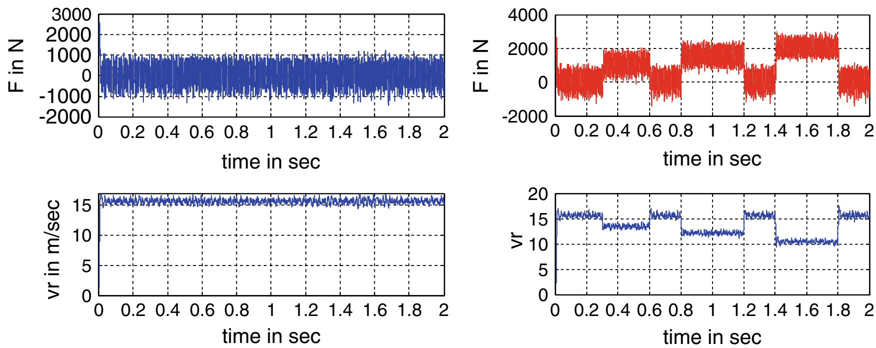
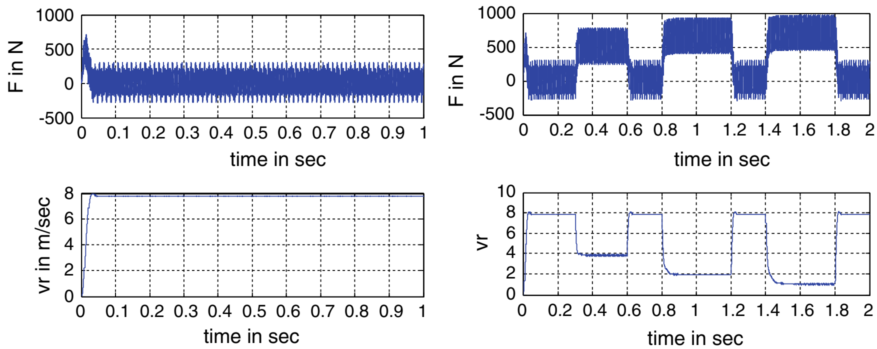


Fig. 10 Effect of changing slip (S) on the $\eta \cos \varphi$



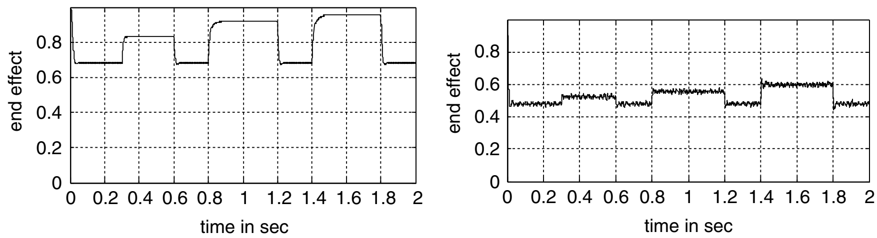
a) At No Load **b)** At variable load

Fig. 11 Force and speed v_r (modulation index = 1)



a) At No Load **b)** At variable load

Fig. 12 Force and speed v_r (modulation index = 0.5)



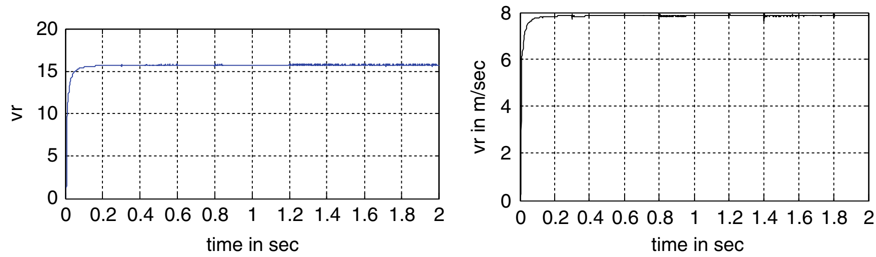
a) At modulation index = 0.5 b) At modulation index = 1

Fig. 13 End effect with time

conditions. Also, it is clear from these results that a wide variable range in speed can be obtained by varying modulation index (*v*/*f* ratio) where when the modulation index equal to 1 the speed equals to the rated speed, and when the modulation index equals to 0.5, the speed reduced to half the rated speed, and the end effect factor is increased due to reduction in speed as shown in Fig. 13. At the instant of load application, there is an increase in force and a reduction in speed; these effects increased with increasing load.

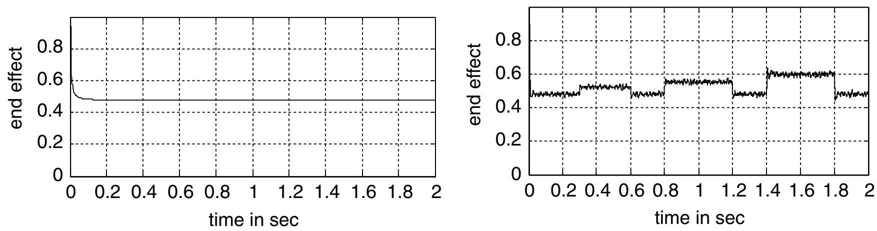
7.2 Simulation Results Using PID Controller and Comparing with VSI

Figures 14, 15, and 16 show the speed and end effect during load intervals with modulation index equal to 1 and at equal to 0.5. If a comparison is made between the speed response without using of PID controller and with using PID controller, it is clear that at the intervals of applied load, the speed decreases without PID controller, while with PID controller, the speed maintains still at the steady-state



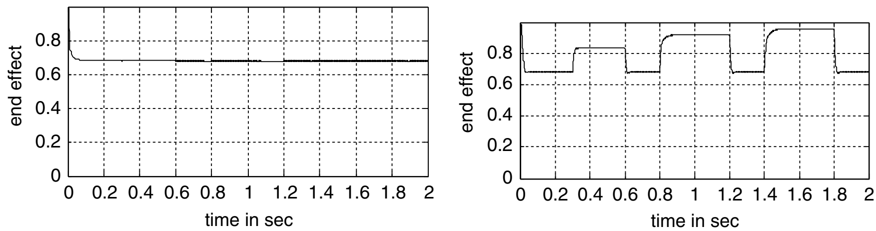
(a) Modulation index = 1 (b) Modulation index = 0.5

Fig. 14 Speed response of SLIM at variable load with PID controller



(a) With PID controller (b) With VSI

Fig. 15 End effect of LIM at modulation index = 1



(**a**) With PID controller (**b**) With VSI

Fig. 16 End effect of LIM at modulation index = 0.5

speed value at the load intervals. This is the requirement of using PID controller for the speed control. Also, using of PID controller will reduce the steady-state error, and it eliminates the overshooting of the speed response with reasonable settling time.

7.3 Result Analysis

After checking the methods and analyzing the result, which is applied to the proposed model, it becomes clear that the power factor can be raised by 12%. Moreover, the efficiency can also be improved by 15%, compared to particle swarm optimization (PSO) method as well as the power factor can be raised by 15%, and the efficiency can be improved by 20% when compared to Cuckoo search. Also, it is noted that increasing the number of poles causes large influence on the thrust of LIM (Velocity vs Force). Moreover, this influence decreases in the ends. However, the loss at end effect divides on the number of poles which means this loss reduces with increasing poles numbers. As a result, the response of the motor becomes better with high number of poles. Similarly, the thrust enhances with increasing poles numbers too. But, experimental outcomes show that increasing poles numbers can lead to shrink the efficiency of the machine. Therefore, the number of poles should be selected depending on the priority ratio between efficiency as well as thrust.

8 Conclusion

In conclusion, this paper gives a simple optimization technique which focuses mainly in enhancing the efficiency and power factor of SLIM. In the beginning, it gives a clear introduction which discuss the role of LIMs in transferring systems, processing materials, smart applications. After that it speaks about the properties of SLIM and the parameters with their limitations which should be applied in the proposed model. After that it demonstrates a brief structure with electrical circuit of the suggested design. Then, a clear theory has been given on the factors which are affecting mainly of the thrust and efficiency as well as power factor. Subsequently, the performance of SLIM is discussed briefly by using simple flowcharts to demonstrate the steps of proposed approaches. Next, the results of all approaches are shown clearly in separate curves. Moreover, all curves and results have been analyzed. Finally, the notes, which are obtained from this analysis, are given in several points. The result illustrates clearly that the power factor can be increased by 15% after applying this technique. Furthermore, the efficiency can be improved by 0.20% by this method, this suggested as one of the simple and cheap methods which can be used to enhance the performance of SLIM for different electrical applications.

References

1. Xu W, Xiao X, Du G, Hu D, Zou J (2020) Comprehensive efficiency optimization of linear induction motors for Urban Transit. IEEE Trans Veh Technol 69(1)
2. Shiri A, Shoulaie A (2012) Design optimization and analysis of single-sided linear induction motor, considering all phenomena. IEEE Trans Energy Conv 27(2)
3. Mishima T, Hiraoka M, Hiraoka, Nomura T (2005) A study of the optimum stator winding arrangement of LIM in maglev systems. In: IEEE international conference on electrical machines and drives, pp 1238–1242, May 2005
4. Higuchi T, Nonaka S (2001) Design of high efficiency linear induction motors for linear metro. Electr Eng Japan 137(2):36–43

5. Isfahani HA, Ebrahimi BM, Lesani H (2008) Design optimization of a low speed single-sided linear induction motor for improved efficiency and power factor. IEEE Trans Magn 44(2):266–272
6. Pourmoosa AA, Mirsalim M (2013) Equivalent circuit of linear induction motor based on coupled circuit model and optimization design using imperialist competitive algorithm. In: Proceedings of the 4th annual international power electronic and drive systems and technologies conference, Tehran, pp 349–354
7. Isfahani A, Hassanpour, Lesani H, Ebrahimi BM (2007) Design optimization of linear induction motor for improved efficiency and power factor. In: Proceedings of the IEEE international conference on electrical machine and drives, May 2007, pp 988–991
8. Hassanpour A, Isfahani B, Ebrahimi M, Lesani H (2008) Design optimization of a low-speed single-sided linear induction motor for improved efficiency and power factor. IEEE Trans Magn 44(2):266–272
9. Bazghaleh A, Naghashan Z, Naghashan M, Meshkatoddini R (2010) Optimum design of single-sided linear induction motors for improved motor performance. IEEE Trans Magn 46(11):3939–3947
10. Boldea I, Nasar SA (2001) Linear electromagnetic device, Taylor & Francis 2001. Wells JR, et al (2001) Linear induction machine design for industrial laboratory development. In: Proceedings of electrical\electronic insulation conference, 2001, pp 319–322
11. Ho D, et al (1993) Design of single side linear induction motor using the finite element method and SUMT. IEEE Trans Magn 29(2):1762–1766
12. Korkmaz F (2014) Simulink model of vector controlled linear induction motor with end effect for electromagnetic launcher system. elektronika ir elektrotechnika 20(1). ISSN 1392-1215
13. Yoon R, et al (1997) Analysis and optimal design of the slit type low speed linear induction motors. In: IEEE international electric machines and drives conference, p TB2-8.1–3
14. Mishima T, Hiraoka M, Nomura T (2005) A study of the optimum stator winding arrangement of LIM in maglev systems. In: Proceedings of the IEEE international conference on electrical machines drives IEMDC, pp 1231–42
15. Bazghaleh AZ, Naghashan MR, Meshkatoddini MR (2010) Optimum design of single-sided linear induction motors for improved motor performance. IEEE Trans Magn 46(11):3939–3947
16. Zayandehroodi H, Nasrabadian A, Anoosheh R (2015) Cuckoo optimization algorithm based design for low-speed linear induction motor. Özel Sayı Sci J 36(6):30–37
17. Lucas C, Nasiri Z, Tootoonchian F (2010) Application of an imperialist competitive algorithm to the design of a linear induction motor. Energy Convers Manag 51(7):1407–1411
18. Creppe RC, Ulson JA, Rodrigues C (2008) Influence of design parameters on linear induction motor end effect. IEEE Trans Energy Convers 23(2):358–362
19. Lu J, Ma W (2011) Research on end effect of linear induction machine for high-speed industrial transportation. IEEE Trans Plasma Sci 39(1):116–120
20. Yang T, Zhou L, Li L (2011) Influence of design parameters on end effect in long primary double-sided linear induction motor. IEEE Trans Plasma Sci 39(1):192–197
21. Xu W, Zhu JG, Zhang Y, Li Y, Wang Y, Guo Y (2010) An improved equivalent circuit model of a single-sided linear induction motor. IEEE Trans Veh Techol 59(5):2277–2289
22. Sharma S, Singh J, Obaid AJ, Patyal V (2021) Tool-condition monitoring in turning process of Fe-0.75Mn-0.51C steel with coated metal carbide inserts using multi-Sensor fusion strategy: a statistical analysis based ingenious approach. J Green Eng 2998–3013

An Approach for Potato Yield Prediction Using Machine Learning Regression Algorithms

Prabhu Prasad Patnaik and Neelamadhab Padhy

Abstract Agriculture is backbone of any country's economy, and also, good crop yield is highly essential for supporting the growing demand of increasing population. By using machine learning, we will be able to predict the crop yield and also the right crop that can be grown in a particular area by analyzing the soil data and the weather data of the particular location. This study mainly focuses on how supervised and unsupervised machine learning approach help in the prediction. Different machine learning algorithms include KNN algorithm, SVM, linear regression, logistic regression, NB, LDA, and decision trees. Taking different dataset preprocessing operation is performed, and missing data are modified so that it does not affect the prediction. Then, the processed data are utilized by the machine learning algorithms for making the prediction. The dataset is divided into training set and test set, and the accuracy of prediction is verified. There are different performance metrics which can be used to evaluate the accuracy in prediction of the algorithms like MSE, MAE, and RMSE, coefficients of determination metrics (R^2), confusion matrix, accuracy, precision, recall, and F1-score.

Keywords Machine learning · Potato yield prediction · Linear regression algorithm

1 Introduction

With the growing population and increasing food demand, agriculture has become an important part of any country and also the chief source of national income. A sound agriculture sector insures a nation of food security. Most countries lean on agricultural products allied industries for their principal source of income. There are many crops that can be grown in a particular area depending upon the climatic condition and soil health of that particular location. Earlier yield prediction was dependent upon an expert view who was well connected with the particular location, and later on,

P. Prasad Patnaik · N. Padhy (✉)
School of Engineering & Technology, Department of CSE, GIET University, Gunupur 765022, Odisha, India
e-mail: dr.neelamadhab@giet.edu

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_27

various traditional methods were implemented to predict crop prediction, but due to rapid urbanization, there had been a great change in the climatic condition of any region due to which the prediction was not accurate which lead to a loss to the farmer monetarily. This also affected the crop selection, i.e., farmers were having less knowledge about the type of crop that they can grow considering the current climatic condition and soil health in that particular region; as a result, the yield was not up to the mark as expected. So, there must be some mechanism which would outperform traditional methods in prediction crop yield and selection of the right crop. Here, ML plays a vital part in making right decision in predicting the right crop and also predicting the yield which would help a farmer in reducing the loss incurred and to have prior knowledge of what crops can be grown in a particular location and what will be the monetary gain from the crop yield and also fulfill the need of the rising population and also prove that agriculture is the real backbone of any country. Machine learning is an important tool which helps in predicting crop yield and crop selection in a particular location. It is a subfield of artificial intelligence (AI), and its main goal is to fit the data into models that can be understood, and people can utilize it. There are many supervised and unsupervised machine learning approach that can be implemented for making the prediction, and the dataset required for implementing machine learning can be obtained from many governments Web site or from the Agricultural Department of any particular location, and also, satellite images can be used. The types of data that can be collected include the weather parameters, soil parameters, and these data are first preprocessed, and then, missing data are modified, and feature selection algorithm is implemented to look for the best feature which would help in making the prediction accurate. The accuracy of any algorithm can be evaluated by using MAE, MAPE, RMSE, and R2. So, machine learning can be used for large-scale yield forecasting or can be used for a small location.

1.1 Critical Contribution

We have taken the help of linear regression algorithms like the SLR, MLR, and LR to predict the potato yield prediction, and for this, we have taken the dataset consisting of yield data of different districts of Andhra Pradesh for the period 2000–2012.

2 Literature Reviews

Schwalbert. et al. [1] proposed three algorithms, i.e., multivariate OLS linear regression, RF, and LSTM, NN to predict soybean yield in southern Brazil and compared their performance by taking satellite information from NASAEOSDIS, and weather data through GEE using the CAR layer. The outcome of the study was LSTM neural net-works showed a lower value of MAE, MSE, and RMSE, and least performance was of OLS. The yield forecast can be more accurate with addition of multi-temporal satellite images along with weather data. Nevavuori et al. [2] proposed convolutional neural networks (CNNs) to predict wheat and barley yield from NDVI and RGB data captured by unmanned aerial vehicles (UAVs) or drones. By using Ada delta training algorithm and CNN with convolution layers, MAE of 484 kg/ha and MAPE of 8.8% were obtained. In the future, climate and soil features can be added along with the image data for better accuracy of prediction. Gopal et al. [3] came up with a hybrid MLR–ANN model for predicting the paddy crop yield and compared it with the finding of ANN alone and found that the hybrid model gave a better accuracy in prediction than the ANN model having an error rate of 0.041 as compared to 0.064 for ANN, and the hybrid approach has high R2. For this study, data were obtained from the Statistical, Meteorological, and Agricultural Departments. In the future, new hybrid models can be implemented and compared with conventional models. Gaso et al. [4] proposed a study for predicting wheat yield by coupling remote sensing data and crop models at field scale. The model used vegetation index data obtained from remote sensing data by Landsat and used simple regression model (SRM) and crop model method (CMM), and observed that using CMM along with satellite data showed high precision as regards SRM on single image, and also, CMM showed better spatial variability on estimated yield. In the future, other machine learning models can be tested for accuracy in yield prediction.

Gopal et al. [5] proposed a study of predicting crop yield using four machine learning algorithms, i.e., ANN, SVR, KNN, and RF and evaluated performance. The data for this study are obtained by feature selection algorithms, i.e., FFS algorithm, CBFS algorithm, VIF algorithm, and RFVarImp algorithm. It was observed that RF algorithm using FFS outperformed from the other three ML algorithms. In the future, other feature sets can be taken, and the prediction accuracy can be compared. Shahhosseini et al. [6] showed how machine learning algorithms, i.e., LASSO regression, ridge regression, RF, XGBoost, and as a whole, can be used as meta-models in predicting maize yield and nitrogen loss from preseason weather information from Daymet and soil information from SSURGO database. Filtering and testing were done by R-Statistical software. XGBoost showed a promising performance in prediction accuracy of RRMSE of 13.4%. In the future, machine learning algorithms can be implemented on different datasets, and the prediction accuracy can be measured. Thomas et al. [7] suggested a model for predicting beforehand which type of crops that can be grown in a specific field depending on the outcome of the model. The approach uses four machine learning algorithms, i.e., KNN, decision tree, KNN with cross-validation, Naïve Bayes, and SVM. The data for this study include the soil properties nitrogen, phosphorus, potassium, and the pH value. After the test, it was found that KNN with cross-validation gave an accuracy of 88% and thus can be used for prediction. In the future, IOT-based implementation can be done for more accurate prediction. Nischitha et al. [8] designed a system to predict crop that can be grown in a particular land and the amount of nutrient that need to be added to the soil for proper yield. The system used SVM algorithm and decision tree algorithm for prediction by taking the dataset containing oil pH, temperature, humidity, rainfall, crop data, NPK values collected from VC Form Mandya, government Web sites, and weather department thus helping the farmers in making right decision. The prediction can also be done from data by giving GPS location and from rain forecasting system of the government.

Ravi et al. [9] proposed an IOT and ML-based yield prediction and soil monitoring system which used SVM algorithm for prediction. The data are gathered from the cloud database server uploaded by node MCU. A mobile application is designed which shows the expected yield of a particular area by taking into account the soil parameters of the particular area. In the future, information quality can be improved for better prediction. Rao et al. [10] came up with a model for forecasting crop yield of a particular area by using supervised machine learning algorithm, i.e., KNN algorithm. The dataset for this model includes the soil dataset, rainfall-related dataset, and crop yield data which were collected from Kaggle.com and data.gov.in. The model predicted accuracy above 75%. In the future, climate data can be added to make the prediction more accurate. Suruliandi et al. [11] proposed a study where different feature selection techniques like RFE, Brutal, and SFFS are applied on the dataset to find out the best attributes, and then, the reduced dataset has been applied to different classifiers like KNN, NB, DT, SVM, RF, and bagging. The dataset includes the characteristics of soil and factors related to environment gathered from www.tnau.ac.in and the Agricultural Department of Sankarankovil Taluk, Tenkasi District, Tamil Nadu, India. In this study, it was that RFE with the bagging method provides better accuracy. Paudel et al. [12] outlined an ML workflow for proving that applying machine learning to forecast crop yield is the right choice. Here, the yield prediction of soft wheat, spring barley, sunflower, sugar beet, potatoes was done in Netherland, Germany, and France. Supervised learning approach was implemented to anticipate early season and end of season yield. The prediction was compared with the prediction of the European Commission's MCYFS, and found that using machine learning to large scale was accurate. In the future, new data sources can be added for more accurate prediction. Suganya et al. [13] proposed a study where crop yield prediction is tested by using supervised machine learning algorithms, i.e., logistic regression, decision tree, random forests, KNN, SVM, and the performance of these algorithms is evaluated by considering precision, recall, and F1-score. The dataset for this study includes soil data, weather data, and past year production data. Based on the criteria of evaluating the ML algorithm, logistic regression gave the highest accuracy of 100%. In the future, other evaluation parameters can be considered for predicting the accuracy.

Mupangwa et al. [14] proposed a study of predicting maize yield under conservations agriculture in Africa using machine learning. The data collected from different countries and multisite were classified into highland and lowland and supervised linear algorithms, i.e., LR and LDA and nonlinear algorithms, i.e., KNN, CART, NB, and SVM were used. To measure the accuracy of the algorithm, accuracy statistics and Cohen kappa statistic were used, and observed that LDA outperformed all other algorithms in accuracy. In the future, the performance can be compared with modeling tools. Palanivel et al. [15] proposed a study where different ML algorithms were implemented for crop yield prediction. Here, linear regression, artificial neural networks, and SVM are being implemented, and the data for this study include soil data and weather data. The accuracy of the algorithms was measured by metrics like MAE, RMSE, and MSE. It was observed that ANN and SVM are well suited for accurate prediction. To take the study further, conceptual approach may be implemented.

3 Data Preprocessing

In data preprocessing, the first step carried out is importing the libraries and then the dataset which is a.csv file's uploaded into the IDE; if the dataset has some missing values, then we can fill it with the mean or the average of all the data in that particular column; the next step is to encode the categorical data; after that the data are divided into training and test set, and the feature scaling is performed. When sending your final files, please include a readme informing the contact volume editor which of your names is/are your first name(s) and which is/are your family name(s). This is particularly important for Spanish and Chinese names. Authors are listed alphabetically according to their surnames in the author index.

4 Proposed Model

The proposed model consists of the input dataset which is first preprocessed where the raw data are prepared to make it suitable for a machine learning model. The processed data are then given for feature selection where the features that are required is selected for its use in the model on which regression algorithm is being applied for making the prediction. The abovementioned Fig. 1 is the proposed model for forecasting crop yield.

Phase-1: In the first phase, we have collected the potato yield data of different districts of Andhra Pradesh.

Fig. 1 Proposed model

Phase-2: In this phase, we have done data preprocessing where the missed data were being replaced by the mean, and also, we have selected the best features from the set of features available in the dataset which would help in making correct prediction.

Phase-3: In this third phase, on the processed data, we have applied regression algorithms, i.e., SLR, MLR, LR, and we have used different performance metrics to measure the accuracy of our prediction.

Phase-4: In this final phase, we have come up with the result of our prediction.

5 Result and Experimental Observations

We had performed the prediction on potato yield dataset of different locations of Andhra Pradesh in different years. We performed SLR, MLR, and logistic regression and compared result of three. SLR is used to find the relationship between one independent variable and dependent variable. The SLR is given by

$$y = a + bx \tag{1}$$

where b represents slope of the line, a represents y-intercept of the line, y represents the dependent variable, and x represents the independent variable.

MLR is used to find the relationship between two or more independent variables and a dependent variable. Mathematically, it can be represented as

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon \tag{2}$$

where Y represents dependent/output variable, X terms are independent variables, β_0 represents y-intercept, β terms represent slope coefficients, ε represents the models error term or residuals.

Logistic regression is a supervised learning algorithm which is used when dependent variable is categorical and uses sigmoid function to model the data and can be represented as

$$f(X) = \frac{1}{1 + e^{-x}}, \tag{3}$$

where $f(x)$ is output between 0 and 1 value and x is input function.

The steps we followed for our study are in the first step we imported the libraries, and then, we imported the dataset into our IDE, and then, we extracted the best feature for our dataset and then performed data preprocessing, i.e., splitting the dependent and independent variables, processing the missing data, splitting the data into training and test set. Then, on to the processed data, we applied the model and measured the performance on the basis of different metrics and found that in case of simple linear regression the R2 score as 0.0222, MSE as 13.461, RMSE as 3.68; for multiple linear regression, the R2 is found to be 0.1572, MSE as 11.602, and RMSE as 3.406, for logistic regression and found the R2 as −0.100078, MSE as 381.5384, and RMSE as 19.533. The below mentioned Figs. 2, 3, and 4 are meant for linear regression as well as multiple linear regression. Figures 5 and 6 are meant for heat map representation of linear and multiple linear regression algorithm.

By taking the confusion matrix as shown in Fig. 7, we can judge the performance of a classification model where it shows the relationship between the actual value and the predicted value. To judge the performance, we have implemented Python to visualize the confusion matrix where the dataset that we have taken is the potato yield data of different districts of Andhra Pradesh.

Fig. 2 Linear regression

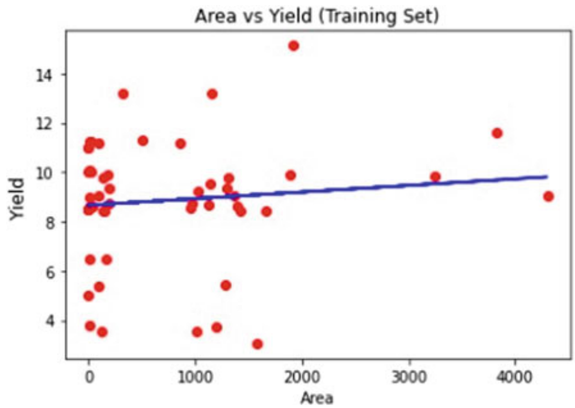
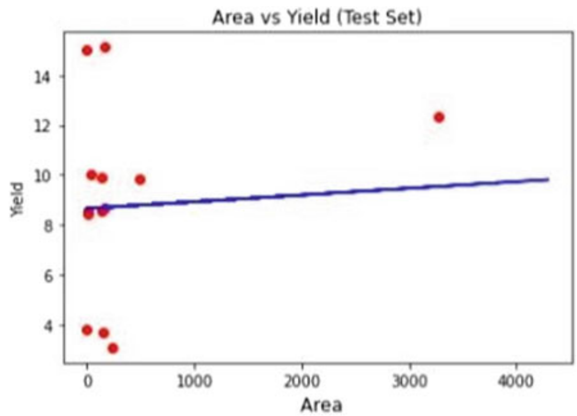


Fig. 3 Multiple linear regression



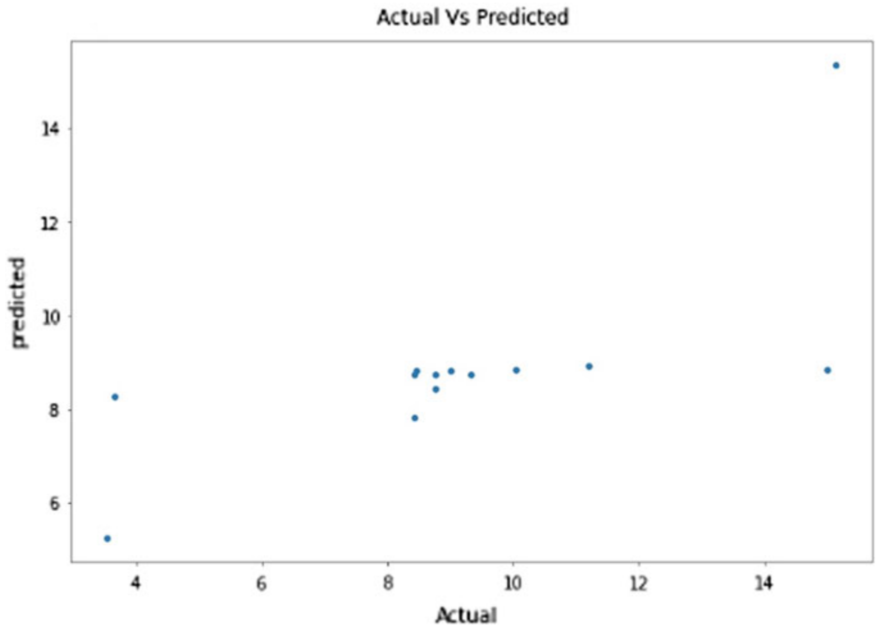
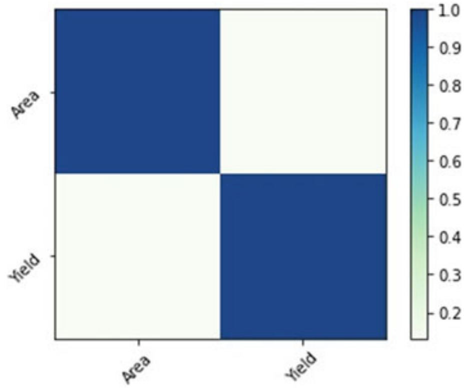


Fig. 4 Multiple linear regression

Fig. 5 Heat map of simple linear regression



6 Conclusion and Future Scope

This study is mainly concerned with the analysis of the role of different ML algorithms in forecasting crop yield. It was observed that we can apply different ML algorithms for the prediction and the selection of algorithm to be used depends on the type of data available based on the location. Different algorithms showed different amount of accuracy taking into account different measuring parameters. We can also

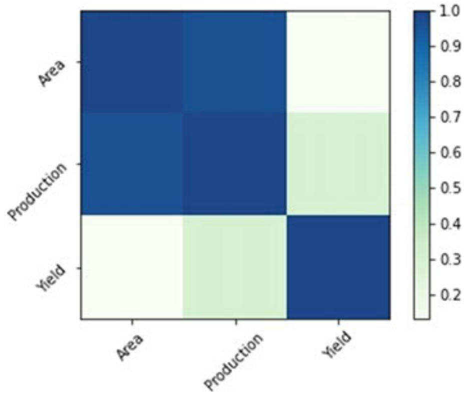


Fig. 6 Heat map of multiple linear regression

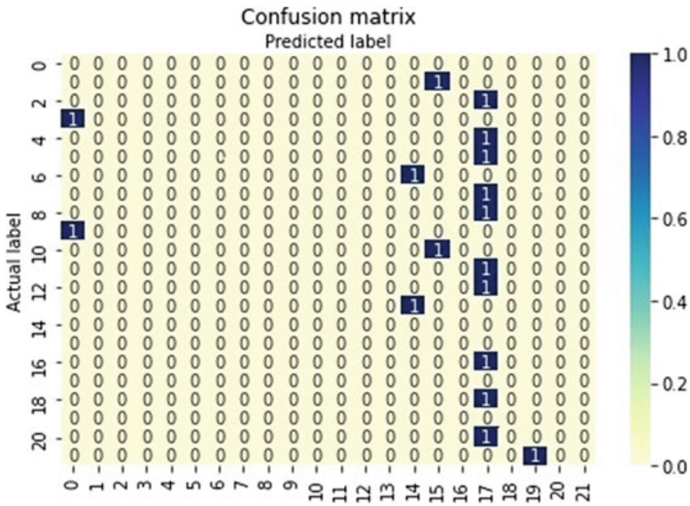


Fig. 7 Confusion matrix for logistic regression

go for hybrid models which will bring about more accuracy in prediction. The data to be used in the dataset should be error free for getting accurate prediction. Thus, machine learning can be used for yield prediction in small scale as well as for large scale. In the future, we can go for other performance metrics to measure the performance of the machine learning algorithm so that we will be able to make more accurate prediction.

References

1. Prasad PV, Schwalbert RA, Amado T, Corassa G, Pott LP, Ciampitti IA (2020) Satellite-based soybean yield forecast: integrating machine learning and weather data for improving crop yield prediction in southern Brazil
2. Nevavuori P, Narra N, Lipping T (2019) Crop yield prediction with deep convolutional neural networks. Comput Electron Agric 163:104859
3. Gopal PM, Bhargavi R (2019) A novel approach for efficient crop yield prediction. Comput Electron Agric 165:104968
4. Gaso DV, Berger AG, Ciganda VS (2019) Predicting wheat grain yield and spatial variability at field scale using a simple regression or a crop model in conjunction with Landsat images. Comput Electron Agric 159:75–83
5. PS MG (2019) Performance evaluation of best feature subsets for crop yield prediction using machine learning algorithms. Appl Artif Intell 33(7):621–642
6. Shahhosseini M, Martinez-Feria RA, Hu G, Archontoulis SV (2019) Maize yield and nitrate loss prediction with machine learning algorithms. Environ Res Lett 14(12):124026
7. Kalimuthu M, Vaishnavi P, Kishore M (2020) Crop prediction using machine learning. In: 2020 third international conference on smart systems and inventive technology (ICSSIT), pp 926–932. IEEE
8. Mahendra N, Vishwakarma hanush, Nischitha K, Ashwini, Manjuraju MR (2020) Crop prediction using machine learning approaches. Int J Eng Res Technol (IJERT) 09(08)
9. Ravi G (2021) Soil monitoring and crop yield prediction using machine learning. Turkish J Comput Mathe Educ (TURCOMAT) 12(11):5498–5503
10. Rao KG, Professor A, Yashwanth K, Goud MS (2021) Crop yield prediction by using machine learning techniques. Ann Rom Soc Cell Biol 25(6): 2551–2557. Accessed 16 Jul 2021
11. Suruliandi A, Mariammal G, Raja SP (2021) Crop prediction based on soil and environmental characteristics using feature selection techniques. Math Comput Model Dyn Syst 27(1):117–140
12. Paudel D, Boogaard H, de Wit A, Janssen S, Osinga S, Pylianidis C, Athanasiadis IN (2021) Machine learning for large-scale crop yield forecasting. Agric Syst 187:103016
13. Suganya M (2020) Crop yield prediction using supervised learning techniques. Int J Comput Eng Technol 11(2)
14. Mupangwa W, Chipindu L, Nyagumbo I, Mkuhlani S, Sisito G (2020) Evaluating machine learning algorithms for predicting maize yield under conservation agriculture in Eastern and Southern Africa. SN Appl Sci 2(5):1–14
15. Palanivel K, Surianarayanan C (2019) An approach for prediction of crop yield using machine learning and big data techniques. Int J Comput Eng Technol 10(3):110–118

Design and Implementing Smart Portable Device for Blind Persons

Rana Jawad Ghali and Karim Q. Hussein

Abstract Advances in mobile technology have made significant improvements in helping the visually impaired. The glasses project helps blind people detect and recognize objects in their environment that they see through a small camera attached to their glasses. This method helps a blind person establish a connection with a nearby object by sending a voice message to an earpiece worn over the blind ear. The goal is to develop intelligent systems that can mimic the human eye. To do this, we use a small device called the "Raspberry Pi 4" that works in a way similar to the human brain, using a camera. It is known as a convolutional neural network algorithm for object recognition using deep learning algorithms. Finally, the moment the features of the image are recognized, the sound of each object is transmitted so that the visually impaired can know about the object in front of them. Python was used to create this project. The results showed that the blind CNN classifier achieved 100 curacy on the COCO dataset.

Keywords Raspberry Pi 4 · Deep learning algorithm · Blind persons and common objects in context · Convolutional neural network

1 Introduction

The human eye is the organ that gives us sight, allowing us to observe and learn more about the world around us than our other senses. We use our eyes in almost every activity, including reading, working, watching TV, writing letters, driving a car, and many other activities. Most people would agree that vision is the most valuable emotion [1]. Blind people are often ignorant of the dangers they face on a daily basis. Even in familiar surroundings, they may confront several obstacles when carrying out their everyday activities. In this study, we proposed a convolutional neural network

R. Jawad Ghali (✉) · K. Q. Hussein
Computer Science Department, University of Mustansiriyah, Baghdad, Iraq
e-mail: ranajawad995@gmail.com

K. Q. Hussein
e-mail: karim.q.h@uomustansiriyha.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_28

(CNN)-based intelligent object identification system to provide a smarter and safer lifestyle for the visually impaired [2]. The project's goals are to promote a globally challenging problem in computer vision, such as object recognition in the context of surrounding items, which is something that the blind does on a regular basis. MS COCO is large-scale object identification, segmentation; it provides necessary environmental information using the camera mounted on the glasses of the visually impaired. To apply the needed recognition, a dataset of items acquired from daily scenes is built. Object detection is a method of detecting real objects in digital images such as bicycles, chairs, doors, and tables, often used in blind environments. Based on their geographical regions, Furthermore, the camera is utilized to detect any things. The suggested approach for the blind seeks to provide persons with visual loss with more opportunities to reach their full potential. The project's main goal is to create and construct a real-time object identification system utilizing blind glass [1]. Creating a tool for a blind person, this is not a new challenge. However, creating a computer-aided tool is still a work in progress. The goal of all of these systems is to assist the blind persons to be a self-relevant. Several works employ computer vision methods. However, there is no known solution that can assist in meeting all of a blind person's fundamental needs [3].

2 Related Work

Here are some examples of related work that uses various approaches to aid blind people with their activities:

- Iswahyudi1, Khairul Anam1, Bambang Sujanarko1, "Development of visual aids for blind people based on faster RCNN." The purpose of this study is to develop a visual aid for the visually impaired that recognizes objects using a faster RCNN. A faster RCNN is installed on a Raspberry Pi with a camera. The result is an audio signal that describes the object and its position with respect to the user [4].
- Adam Gunnarsson, "Real time object detection on a Raspberry Pi." The Raspberry Pi 3's suitability for conducting object detection, a popular embedded computer board, is investigated in this thesis. Two controlled tests are carried out in which two cutting-edge object detection models, SSD and YOLO, are put to the test in terms of speed and accuracy. In YOLO, 4-m accuracy is 67 percent, whereas SSD accuracy is 93% [5].
- Ashwani Kumar, S. S. Sai Satyanarayana Reddy, and Vivek Kulkarni, Real-time object detection techniques for the visually impaired using deep neural networks. In this study, we proposed a real-time object identification method that allows blind people to recognize objects on any device running the model. A single multi-block detector method was used using the architecture of a faster area convolutional neural network. The standard datasets VOC and COCO are used. The model is 75 times more accurate [6].

- Y. C. Wong, J. A. Lai, S. S. S. Ranjit, A. R. Syafeeza, N. A. Hamid, "Convolutional Neural Networks for Object Detection Systems for the Visually Impaired." In this study, we proposed an intelligent object recognition system based on a convolutional neural network (CNN) and created an intelligent and safe living environment for the visually impaired with Raspberry Pi. SSDs have the highest average accuracy of 73.7% compared to faster RCNN 73.2% and YOLO 63.4% [7].
- Mohammad ALNajjar, Ihab Suliman, Ghazi AlHanini, "Real-time Object Recognition and Visually Impaired Recognition." The Smart Spectacles project attaches a small camera to the glasses so that visually impaired or visually impaired people can see and recognize office items. Use a neural network on the Raspberry Pi 3 [1].

3 Phase One—Preprocessing Images

This task uses Microsoft COCO to recognize the object using Per-instance segmentation to support accurate localization of the object. Images of 91 distinct object types that a four year old may recognize are included in our collection. There are 2.5 million labeled occurrences in 328 k photos for a total of 2.5 million tagged occurrences [8]. Divide the result correction by 30% of the test data and 70% of the training data to get the data you need for training and testing purposes:

i. **The Conversion of color image to gray scale**

Converting a color image to a grayscale image is the aim to preserve as much information as possible from the original-colored image. In the event of grayscale conversion, further information about color pictures is required. Each color of a pixel in an image is represented by three colors: red, green, and blue (RGB). To convert a color image to a grayscale image, you need to convert RGB values (24 bits) to grayscale values (8 bits). Equations can be used to convert to gray scale using weighting or luma techniques (1).

$$\mathrm{GR} = 00.56\mathrm{G} + 00.33\mathrm{R} + 00.11\mathrm{B} \tag{1}$$

Green color gives 56%, which is greater than the other colors, according to the abovementioned calculation. The color red provides 33%, followed by the color blue, which contributes 11% [9] (Fig. 1).

ii. **Enhance the image by histogram Equalization**

In digital signal processing, the number of pixels (frequency) having the same intensity value is recorded in an array of histograms, commonly referred to as "bins" (DSPs). Because the intensity range of an 8-bit grayscale picture is 0 to 255, the size of a histogram bin for an 8-bit grayscale image is 256 [10].

We have a digital photograph A that is $M \times N$ in size.

Fig. 1 **a** Colored image **b** grayscale image

$$P(Z) = \frac{\sum_{i=1}^{M \times N} \delta(Z, Zi)}{M \times N} \tag{2}$$

where $\delta(z, y) = \begin{cases} 1, & z = y, \\ 0, & z \neq y, \end{cases}$ and $z \in [z_0, z_L]$,
And (Z) representing the intensity [11] (Fig. 2).

iii. **Use bilinear interpolation to resize images**

A weighted average of the four nearest pixels is used to fill the interpolated points. This method was used to perform two types of linear interpolation: horizontal and vertical. With bilinear interpolation, four interpolation functions must be computed for each grid point. Interpolation kernel for linear interpolation [12]:

Fig. 2 Image enhancement

Fig. 3 Image size (20 × 20)

$$u(y) = \begin{matrix} \{0 & |y| > 1 \\ \{1 - |y| & |y| < 1 \end{matrix}$$

The distance between the interpolated point and the grid point is defined as y.

This is where the image is scaled to 20 × 20 pixels (Fig. 3).

iv. **(PCA) Features extraction**

Two-dimensional PCA (2DPCA) was created to diminish the computing taken a toll of ordinary PCA (too alluded to as picture PCA (IMPCA)). Unlike PCA, which interprets pictures as vectors, 2DPCA considers them to be matrices. 2DPCA produces an eigenvalue issue when a suitable criterion is used although it has a considerably smaller dimensionality than PCA [13].

$$J(x) = X^{\mathrm{T}}\left[\sum_{a=1}^{n} C(a)^{T} C(a)\right]x \tag{3}$$

4 Phase Two—Deep Learning

Convolution neural network (CNN model): CNNs are various leveled neural systems in which convolutional and subsampling layers interchange, taking after basic and complex cells within the essential visual cortex. The conv. and subsampling layers, as well as how the nets are instructed, vary among CNNs [14] (Fig. 4 and Table 1).

5 Phase Three—Raspberry Pi V.4

Computers are being made in a brief period of time, with higher speed, equipment, computer program, lower costs, and expanded get to innovation accessibility. As a result, different endeavors on assistive innovations to permit localization, route, and protest acknowledgment have been set up. The ideal client interface can at that point be changed based on a client's ask, such as vibrations, sounds, or talked words [1].

The Raspberry Pi can run a wide range of software, including a variety of operating systems (the software that allows a computer to function). Raspbian is the most widely

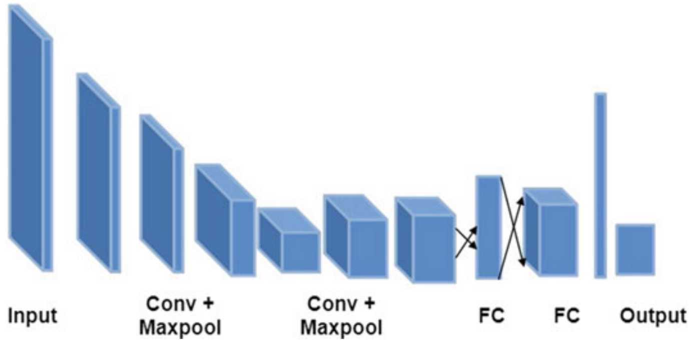


Fig. 4 CNN layers

used of these and is the official operating system of the Raspberry Pi Foundation. Raspbian is based on Debian Linux and is tailored for the Raspberry Pi, with a variety of extras preinstalled and ready to use [15–18] (Fig. 5).

The Raspberry Pi Establishment created an arrangement of reasonable single-board computers within the Joined together Kingdom to empower the instructing of fundamental computer science in schools and creating nations. The introductory show was essentially better known than anticipated, offering for employments such as mechanical technology exterior of its aiming showcase. The Raspberry Pi does not come with any peripherals, such as keyboards, mouse, or cases. However, several accessories have been included in a variety of official and unofficial bundles [1].

Raspbian is the most widely used of these and is the official operating system of the Raspberry Pi Foundation. A few of these components are required, whereas others are discretionary; in any case, all Raspberry Pi models share the same CPU, the BCM2835, which is cheap, effective, and energy-efficient. The Raspberry Pi is comparable to a commonplace PC in that it requires a console for the command section, a show unit, and a control source. The SD Streak memory card, which is commonly utilized in computerized cameras, is set to appear to the Raspberry Pi's processor as a hard drive. The micro-USB port is used to power the device. You can connect to the Internet using an Ethernet/LAN cable or USB dongle (Wi-Fi connection) [9].

i. Main Accessories

In addition to the Raspberry Pi, there are other accessories connected to it to make the process complete (Fig. 6):

(a) **Camera**: Connecting the Raspberry Pi camera last version (v2.3) with the Raspberry Pi 4 microcomputer through flat flexible cable, it attaches thru a 15 cm (Ribbon cable) to the CSI port at the R-Pi.

(b) **Power supply**: Connecting power supply with the Raspberry Pi 4 through USB cable. The power supply ought to give 5 V DC and a least appraised current of 3A. So, the choice was to utilize power bank J34 compelling source of mobile

Table 1 Layers and filter size of CNN model

Ft-s = 16	Ft-s = 32	Ft-s = 64	Ft-s = 128	Ft-s = 256	Ft-s = 512	Ft-s = 1024	Ft-s = 1024	Ft-s = 125	Flatten	Dense
1-D Conv Str = 1 Kr = 3 Max-pool Str = 1 Pool-size = 1	1-D Conv Str = 1 Kr = 3 Max-pool Str = 1 Pool-size = 1	1-D Conv Str = 1 Kr = 3 Max-pool Str = 1 Pool-size = 1	1-D Conv Str = 1 Kr = 3 Max-pool Str = 1 Pool-size = 1	1-D Conv Str = 1 Kr = 3 Max-pool Str = 1 Pool-size = 1	1-D Conv Str = 1 Kr = 3 Max-pool Str = 1 Pool-size = 1	1-D Conv Str = 1 Kr = 3 Max-pool Str = 1 Pool-size = 1	1-D Conv Str = 1 Kr = 3	1-D Conv Str = 1 Kr = 3		Soft-max

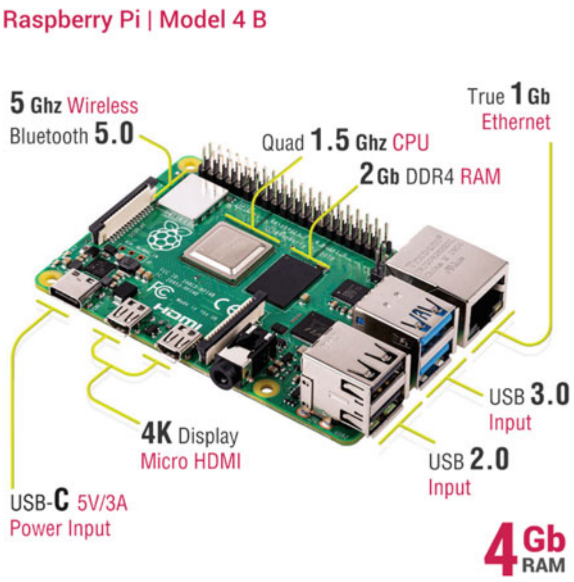


Fig. 5 Structure of Raspberry Pi 4

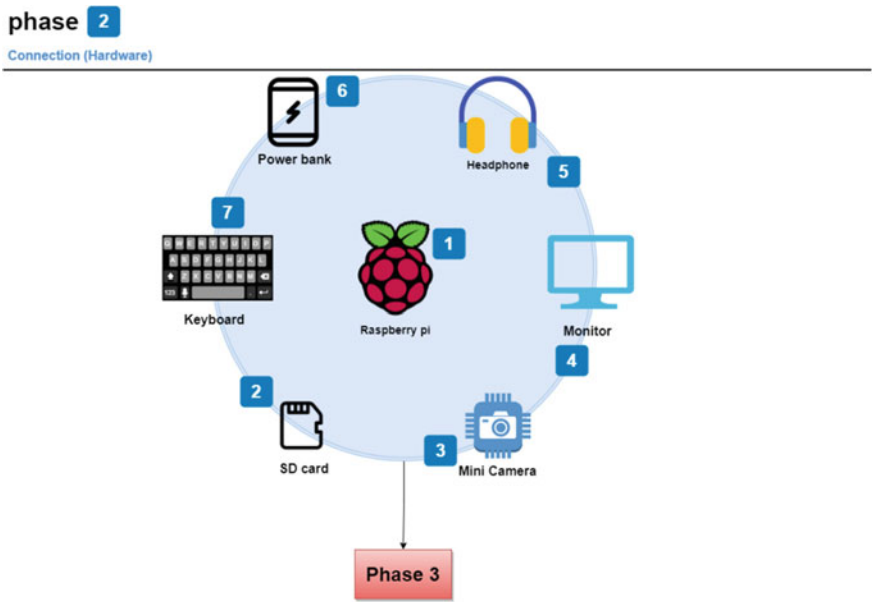
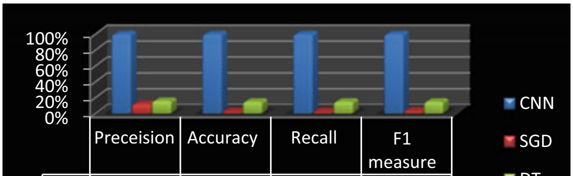


Fig. 6 Raspberry Pi and its accessories

Fig. 7 Result after using different techniques



	Preceision	Accuracy	Recall	F1 measure
CNN	100%	100%	100%	100%
SGD	12%	2%	2%	3%
DT	16%	15%	15%	15%

versatile battery 10000 mAh 2A output numerous circuit assurance fire retardant PC fabric.
(c) **One sunglass**: To fix the camera on it.
(d) **One SD card**: Uploading the OS system for the Raspberry Pi and the proposed system on it.
(e) **One headphone**: Connecting wired headphones to the Raspberry Pi through 3.5-mm audio-out port to send back the name of the objects to the blind person.
(f) **Mini Screen 7 inch HDMI Display-B 800 × 480 pixel**: Used to enter inside the R-Pi. This is connecting to the Raspberry pi through HDMI cable.

6 Results

After working on the dataset COCO, training and testing the results from using different techniques were very uneven. Figure 7 shows the results from using machine learning; we used two algorithms (DT and SGD) when the results were too low, while the results from using deep learning (CNN algorithm) were too high.

While the results using CNNs in deep learning are high (fidelity, precision, recall, and measures F1 = 100%).

While machine learning results that SGD was (precision is 12%; accuracy is 2%; recall is 2%, and F1 measure is 3%).

And DT was (16% precision, 15% precision, 15% recall, 15% F1 measure).

7 Conclusion and Future Research

In this paper, the researcher intends to implement a system in the internal environment so that visually impaired people can use the system on a daily basis to become independent and move around the house, and some work has been done in this field. In the course of this research, we created a CNN model for identifying images of objects. The model was trained on a large dataset. And the accuracy of the model was 100%.

This project is updated by developing models to work in high fidelity dynamic environments and provide additional specifications for the visually impaired. And we will work with AI to try to make devices smarter.

Acknowledgements I'd want to thank Dr. Karim Qasim for his assistance and direction in completing this work on time. Also, thanks to Al Mustansiryha University for their assistance.

References

1. AL-Najjar M, Suliman I, Al-Hanini G (2018) Real time object detection and recognition for blind people
2. Ahmed Ibrahim YAE, Fawzy El-Ghannam WI, Hamada AM, Mohamed El-Shaarawy AH, Hassan Amer AM, Hussein Maghraby AM (2017) Smart glasses for blind people
3. Dharanidharan J, Puviarasi R, Boselin Prabhu SR (2019) Object detection system for blind people
4. Iswahyudi, Anam K, Sujanarko B (2019) Development of a visual aid tool for blind people based on faster R-CNN
5. Gunnarsson A (2019) Real time object detection on a Raspberry Pi
6. Kumar A, Sai Satyanarayana Reddy SS, Kulkarni V (2019) An object detection technique for blind people in real-time using deep neural network
7. Wong YC, Lai JA, Ranjit SSS, Syafeeza AR, Hamid NA (2019) Convolutional neural network for object detection system for blind people
8. Lin TY, Maire M, Belongie S, Ross LB (2015) Microsoft COCO: common objects in context
9. Dhanoon BN, Ali HH (2013) Handwritten Hindi numerals recognition. Int J Innov Appl Stud 3(1)
10. Dorothy R, Joany RM, Joseph Rathish R, Santhana Prabha S, Rajendran S (2015) Image enhancement by Histogram equalization
11. Cheng HD, Shi XJ (2004) A simple and effective histogram equalization approach to image enhancement
12. Prajapati A, Naik S, Mehta S (2012) Evaluation of different image interpolation algorithms
13. Wang L, Wang X, Zhang X, Feng J (2005) The equivalence of two-dimensional PCA to line-based PCA
14. Cires¸an DC, Meier U, Masci J, Gambardella LM, Schmidhuber J, Flexible, high performance convolutional neural networks for image classification
15. Halfacree G (2018) The official Raspberry Pi beginner's guide. Raspberry Pi Trading Ltd., p 240
16. Vujović V, Maksimović M (2014) Raspberry Pi as a wireless sensor node: performances and constraints. In: Proceedings of the 2014 37th international convention on information and communication technology electronics microelectronics MIPRO, 2014, June, pp 1013–1018. doi: https://doi.org/10.1109/MIPRO.2014.6859717
17. Bansal R et al (2021) J Phys Conf Ser 1963:012170
18. Abdulbaqi A, Younis M, Younus Y, Obaid A (2022) A hybrid technique for EEG signals evaluation and classification as a step towards to neurological and cerebral disorders diagnosis. Int J Nonlinear Anal Appl 13(1):773–781. https://doi.org/10.22075/ijnaa.2022.5590

Content-Based Image Retrieval Using Multi-deep Learning Models

Bui Thanh Hung

Abstract CBIR—content-based image retrieval is commonly known as the process or technique to the "image retrieval" problem that is the problem of analyzing and searching for a real content of images. Image search is a search technique that uses images as a source to retrieve an image that is similar to the given image. This technique has many applications in various fields and industries: securities, banking, education, business, etc. For the past decade, a variety of approaches has been introduced to solve image search problem, and one of the approaches which proves to deliver the highest results is the deep learning CNN model. This paper presents the approach of using multi-deep learning algorithms and similarity measurement. The problem is solved with three pre-trained CNN deep learning models: RestNet50, RestNet101, and VGG19 to extract features, then based on these features to calculate the cosine similarity between images to find the mostly similar images with the given query image. We obtain some encouraging results from several experiments on flower dataset. The results show that the CNN method has succeeded in supporting the retrieval task and therefore has huge potential for practical applications.

Keywords Deep learning · Content-based image retrieval · CNN · Cosine similarity

1 Introduction

It is obviously seen that the emergence of digital technology and appearance of many devices used in image acquisition have led to the rapid increase of the image databases stored on the Web. Mary Meeker, an expert in Internet analytics and technology at Cornell University (USA), said in the annual report on Internet trends: "1.8 billion digital images are uploaded in a day and 657 billion photos in a year. This means that every two minutes we take more pictures than we took 150 years ago" (Meeker, 2014). This presents major challenges for the traditional way of organizing and searching

B. Thanh Hung (✉)
Artificial Intelligence Laboratory, Faculty of Information Technology, Ton Duc Thang University, 19 Nguyen Huu Tho Street, Tan Phong Ward, District 7, Ho Chi Minh City, Vietnam
e-mail: buithanhhung@tdtu.edu.vn

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_29

photos. Therefore, the need for new image retrieval techniques has become apparent and urgent. Existing technology typically uses the method through text-based image retrieval to make the search simple and easy. However, this method still has some limitations that make the results inaccurate or not meeting the needs of the user performing the search. For text-based image search, the descriptions attached to the image might not always be accurate since the texts which are used to describe for image labeling do not always reflect clearly what the image represents.

The traditional content-based image search method relies entirely on visual features such as texture, color, shape, and local features extracted from the images. This method still faces the issues of how to identify and select the representative features that have a high influence on the accuracy of the search results. This process will be time-consuming; in addition, the problem arises due to the semantic gap between the high-level and the low-level features. This paper proposes a method to build a content-based image retrieval system based on multi-deep learning models of convolutional neural networks to take advantage in images retrieval by the content. We used convolutional neural network to extract features and calculated the image similarity based on cosine similarity. The cosine distance is larger, the higher of similarity of the two images is. The rest of the paper is constructed as follows: Related works are reviewed in Session 2; our proposed method is presented in Sect. 3; Sect. 4 describes our experiments; Conclusion and future direction are mentioned in Sect. 5.

2 Related Works

In recent years, convolutional neural network—a deep learning method has achieved great success in processing a large database. It has proven to be very effective in the field of image processing such as: automatic image coloring [1], pedestrian detection [2], face detection [3], and image classification [4]. In image retrieval, many researchers used CNNs to solve this problem [5–8]. Kiapour et al. [9] conducted a study to find fashion clothing products by similar images on e-commerce Web site. The authors have performed and compared a number of methods, in which the most prominent is the method using the hidden and experimental two-layer CNN network on the Exact Street2Shop dataset. Search engines such as Google, Bing, or the social network Pinterest have also researched and applied to the image retrieval by using deep neural networks. This research is different with others. We propose multi-deep learning methods to content-based image retrieval basing on pretrained effective CNN models such as Restnet50, Restnet101, and VGG. We extract feature by the model with the highest scores to build image Web search.

3 The Proposed Method

Our proposed model has two main parts: the training and the test model, in which the pretrained CNN deep learning models are the framework for feature extraction. These features will be used to calculate cosine similarity score. Base on those results, we will find the most similarity images for the query image. Figure 1 shows the framework of our proposed model.

3.1 Multi-Deep Learning Models

Convolutional Neural Network

Convolutional neural network (CNN) is built based on multilayer neural network; each layer belongs to one of three types: convolution, subsampling, fully connection as shown in Fig. 2 [10–15].

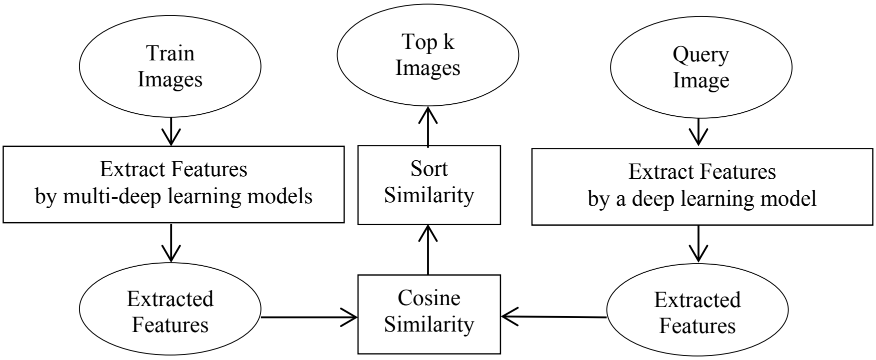


Fig. 1 Proposed model

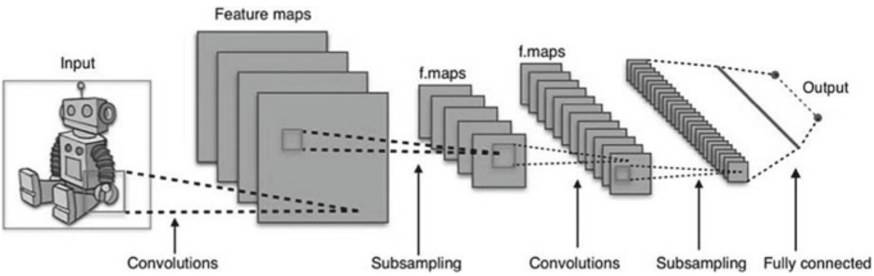


Fig. 2 CNN model

CNN considers the input image as the input layer; each pixel is a neuron; this image is also called as a feature map. Convolution layer acts as a feature extractor that is one or more outputs of the previous layer are convolved with one or more kernels to produce one or more feature maps. Subsampling layer helps the network to withstand data distortions such as translation, rotation, and tilt. Fully connected layer performs classification.

RestNet50, RestNet101

ResNet [12] is a winner model of ILSVRC 2015. This model uses skip connection to fit the input from the previous layer to the next layer and has a deep network of up to 152 layers.

The ResNet50 model has over 23 million trainable parameters which consists of five stages of identity and convolution block. Each identity block has three convolution layers, and convolution block has also three convolution layers.

ResNet101 model has 101 deep layers and can learn rich feature representations.

VGG19

VGG19 [16] is a model built based on convolutional neural network with three fully connected layers and 16 convolutional layers.

3.2 Cosine Similarity

Image search models often look for similarity by category. Two photos are considered similar if they have the same label, e.g., a photo of flowers, a photo of a person, etc. However, this level of similarity is still not sufficient for higher-level applications. The next problem is to find similarity in more details, even when the two images have the same label, in order to identify even the smallest difference.

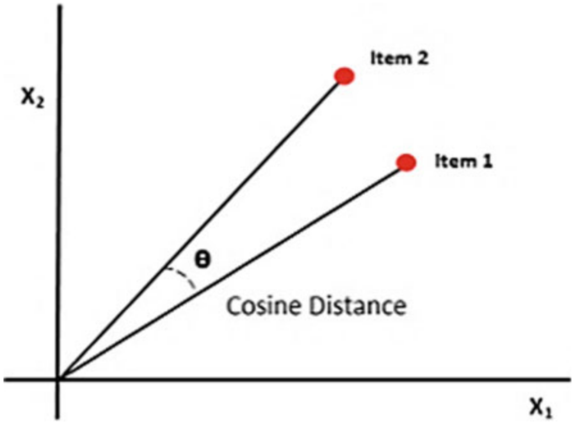
In the past prior to the explosion of deep learning, traditional image similarities based on HOG, SIFT, etc., were commonly used. Nowadays, the CNN, which is often used for image classification and recognition, has presented a more efficient and accurate alternative to the traditional methods.

To compare the image similarity from the obtained feature vectors, we use the cosine similarity measurement. The larger of the cosine distance, the higher the similarity of the two images; based on the distance, it is possible to determine the top k images that are mostly similar to the query image. Figure 3 describes about cosine distance.

Cosine similarity measurement is calculated as follows:

$$\text{Cosine similarity} = \cos(\theta) = \frac{A.B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \tag{1}$$

Fig. 3 Cosine distance



3.3 Content-Based Image Retrieval

With the query image I_q and the image database P, let V_q and V_{iP} be the feature vectors of the query image I_q and the image I_i in the set P, respectively. We determine the degree of similarity between I_q and I_i is the cosine distance between their respective feature vectors extracted by pretrained deep learning models. Finally, the model will use a loss function that is a cross-entropy function defined as follows:

$$L(y, \hat{y}) = -\sum_{j=0}^{M}\sum_{i=0}^{N}\left(y_{ij} * \log\left(\widehat{y_{ij}}\right)\right) \tag{2}$$

The cross-entropy function compares the distribution of the output, with the probability of the correct label being 1 and the probability of the other labels being 0. In other words, the resulting label will be translated into a vector with 1 being the true label; 0 is false label. Then, this vector will be compared with the output of the neuron. The smaller the difference between the two vectors, the smaller the cost function value.

The image retrieval method applied in this paper consists of two stages: feature extraction and similarity comparison with images in the same label to find the images with the highest similarity. All images in the image dataset will be fed through the CNN network to create representative vectors X_i. Then, the representative vector X' of the input image will be compared with each vector X_i obtained above by a cosine similarity measurement, and the mostly similar images to the input image will be returned as the search results. Searching by the method of sorting results images in descending order and shows the top k images most similarity with the query image.

4 Experiments

4.1 Dataset

In this paper, we use the Kaggle flower dataset [17] containing 4242 images of flowers with five classes: sunflower, daisy, rose, tulip, and dandelion with 800 photos of each class. Some pictures are shown in Fig. 4. Table 1 shows detail about the dataset.

Firstly, we did experiment on classification task with three deep learning models: RestNet50, RestNet101, and VGG19. We used ratio 2:8 to divide the dataset to testing and training. Experiments were performed on Google Colab with deep learning library of Keras, TensorFlow, and OpenCV image processing. We evaluated this task by precision, recall, F1-score, and accuracy measurement calculated as follows:

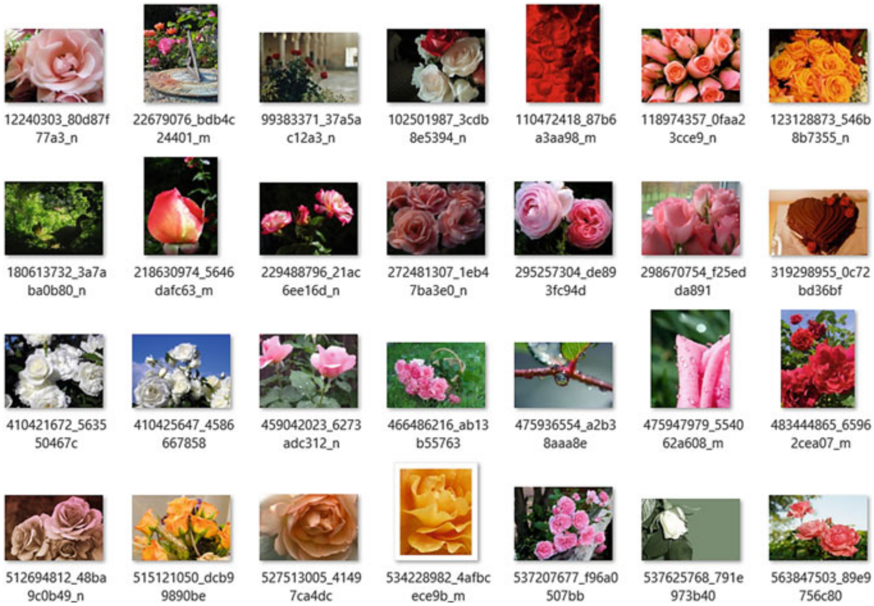


Fig. 4 Photos in Kaggle flower dataset

Table 1 Kaggle flower dataset

Flower	Number
Daisy	769
Tulip	984
Rose	784
Sunflower	734
Dandelion	1055

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

where: *TN*: true negative, *TP*: true positive, *FN*: false negative, *FP*: false positive.

We got the results of RestNet50 model shown in Fig. 5 and Table 2, RestNet101 model shown in Fig. 6 and Table 3, and VGG19 model shown in Fig. 7 and Table 4. The result of three models is shown in Table 5.

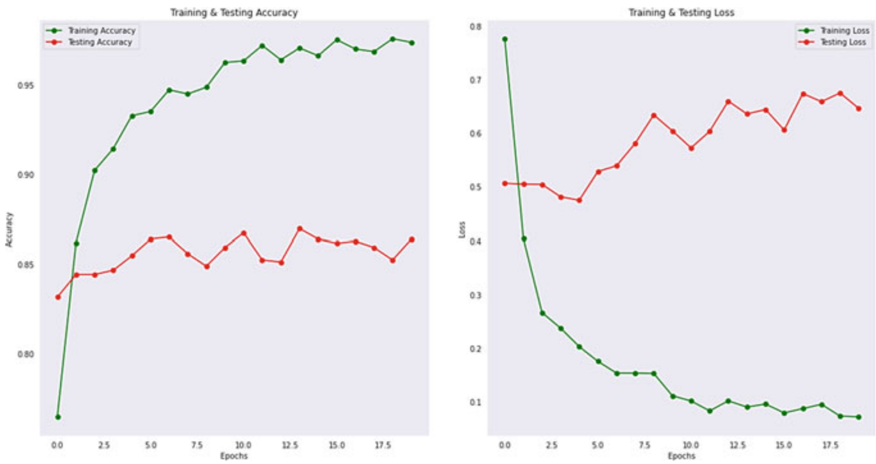


Fig. 5 Result of RestNet50 model in training and testing

Table 2 Results of RestNet50 model

Measurement	Result (%)
Precision	86.6
Recall	86.37
F1-score	**86.4**
Accuracy	**86.37**

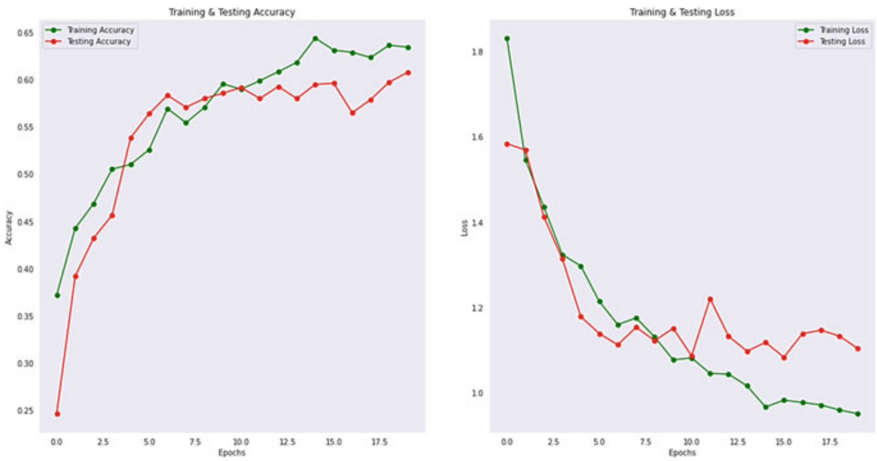


Fig. 6 Result of RestNet101 model in training and testing

Table 3 Results of RestNet101 model

Measurement	Result (%)
Precision	58.99
Recall	58.89
F1-score	**58.63**
Accuracy	**58.89**

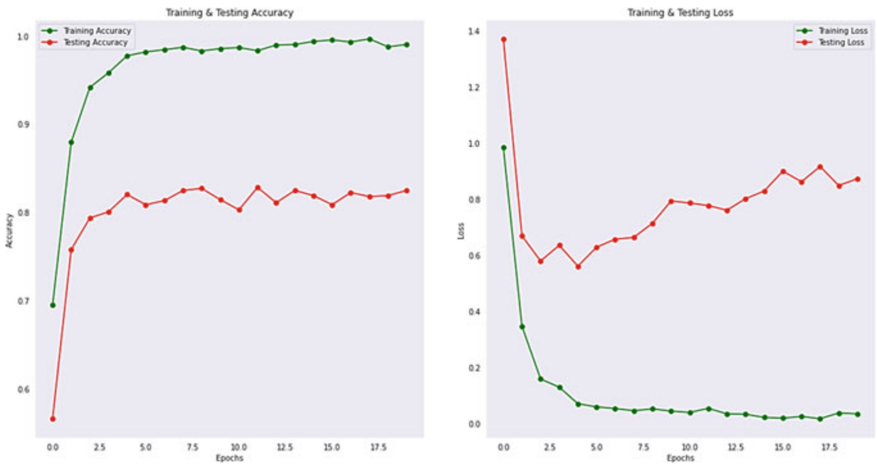


Fig. 7 Result of VGG19 model in training and testing

Table 4 Results of VGG19 model

Measurement	Result (%)
Precision	82.5
Recall	82.45
F1-score	**82.4**
Accuracy	**82.45**

Table 5 Results of three deep learning models

Model	Accuracy (%)
Resnet50	**86.6**
Resnet101	58.99
VGG19	82.5

From Table 5, we saw that RestNet50 is the best model in classification task. We used this model for CBIR task.

Secondly, we did experiment on CBIR based on the RestNet50 model. We extracted feature from this model and calculated cosine similarity to find the top k similar images. The result is shown in Fig. 8.

Fig. 8 Top 6 similar images with the query image (above picture)

5 Conclusion

In this study, we have presented a method to extract features using multi-deep learning model of CNN and integrate these features to calculate cosine similarity to find the most similarity on Kaggle flower dataset consisting of five labels for query image search. The method has the advancement of taking advantage of the processing capabilities of the convolutional neural network for both classification and vector computations representing images in the search image. This study has shown that the application of the RestNet50 model has good results to contribute to improving the efficiency of image search systems. In the future, we will reduce the training time, improve the quality of the image classification, and apply another method to improve the search results.

References

1. Zezhou Cheng, Qingxiong Yang, Bin Sheng: Deep Colorization. ICCV (2015)
2. Wenhan Luo, Junliang Xing, Anton Milan, Xiaoqin Zhang, Wei Liu, Xiaowei Zhao, Tae-Kyun Kim: Multiple Object Tracking: A Literature Review. arXiv:1409.7618 (2014)
3. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich: Going Deeper with Convolutions. In IEEE Conference on Computer Vision and Pattern Recognition- CVPR (2015).
4. Alex Krizhevsky, Ilya Sutskever, and Geo rey E. Hinton: ImageNet Classification with Deep Convolutional Neural Networks. In Proc. NIPS (2017).
5. W. Zhou, H. Li, and Q. Tian: Recent Advance in Content Based Image Retrieval: A Literature Survey. arXiv: 1706.06064 (2017)
6. Benyi Hu, Ren-Jie Song, Xiu-Shen Wei, Yazhou Yao, Xian-Sheng Hua, Yuehu Liu: PyRetri: A PyTorch-based Library for Unsupervised Image Retrieval by Deeps Convolutional Neural Networks (2020).
7. Björn Barz, Kai Schröter, Moritz Münch, Bin Yang, Andrea Unger, Doris Dransch, Joachim Denzler: Enhancing Flood Impact Analysis using Interactive Retrieval of Social Media Images (2019).
8. Zafar B, Ashraf R, Ali N et al (2018) A Novel Discriminating and Relative Global Spatial Image Representation with Applications in CBIR. Appl Sci 8(11):2242
9. M. Hadi Kiapour, Xufeng Han, Svetlana Lazebnik, Alexander C. Berg, and Tamara L. Berg: Where to Buy It: Matching Street Clothing Photos in Online Shops, In Proc. ICCV, (2015).
10. Bui Thanh Hung. Document Classification by Using Hybrid Deep Learning Approach. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering - LNICST, volume 298, pp 167–177 , Springer. (2019)
11. Bui Thanh Hung, Vijay Bhaskar Semwal, Neha Gaud, Vishwanth Bijalwan: Violent Video Detection by Pre-trained Model and CNN-LSTM Approach. Proceedings of Integrated Intelligence Enable Networks and Computing. Springer Series in Algorithms for Intelligent Systems (2021).
12. K. He, X. Zhang, S. Ren and J. Sun: Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition -CVPR, pp. 770–778. (2016)
13. Bui Thanh Hung, Prasun Chakrabarti: Parking Lot Occupancy Detection Using Hybrid Deep Learning CNN-LSTM Approach. International Conference on Artificial Intelligence: Advances and Applications- ICAIAA, (2021).

14. Bui Thanh Hung, Le Minh Tien: Facial Expression Recognition with CNN-LSTM". Research in Intelligent and Computing in Engineering. Springer Series in Advances in Intelligent Systems and Computing (2021).
15. Bui Thanh Hung: Face Recognition Using Hybrid HOG-CNN Approach. Research in Intelligent and Computing in Engineering. Springer Series in Advances in Intelligent Systems and Computing (2021).
16. K. Simonyan and A. Zisserman: Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556, (2014).
17. Alexander Mamaev. (2018) Kaggle Flower Dataset. https://www.kaggle.com/alxmamaev/flowers-recognition.

Design and Simulation of Meander Line Antenna for Operating Frequency at 2.5 GHz Based on Defected Ground Structure

Mohammed Sadiq, Nasri Bin Sulaiman, Maryam Biti Mohd, and Mohd Nizar Hamidon

Abstract Meander line antenna (MLA) with defected ground structure, which resonant at 2.5 GHz, has been designed and examined in this paper. The antenna was built on a FR4 ($r = 4.5$) substrate with a thickness of 1.1 mm and a loss tangent of 0.025. To evaluate the antenna's performance, features were used operational bandwidth, gain, return loss, and radiation pattern. We achieve a return loss of −17 dB, a bandwidth of 57 MHz, and a gain of 3.21 dB using defective ground structure (DGS). The antenna is 34 28 1.1 mm^3, which is a relatively small space.

Keywords S parameter · Antenna · MLA · SAR

1 Introduction

The growth of fast-developing wireless communication applications, it was aided by the development of compact-integrated printed antennas. Because of their advantages of being lightweight, small, conformal, and having a high bandwidth, they are becoming more widely utilized in wireless communication systems [1]. Meander line is the best solution for wireless communication applications such as radio frequency identification tags, USB dongles, Bluetooth headsets, mobile phones, and so on. The

M. Sadiq (✉)
Dhi-Qar Investment Commission, Dhi-Qar, Iraq
e-mail: engmohammedalkaabi@gmail.com

N. B. Sulaiman
Department of Electrical and Electronic Engineering, Faculty of Engineering, University Putra Malaysia, 43400 Serdang, Malaysia
e-mail: nasri_sulaiman@upm.edu.my

M. B. Mohd
Department of Electrical and Electronic Engineering, Universiti Putra Malaysia, Serdang, Malaysia
e-mail: maryam@upm.edu.my

M. N. Hamidon
Electrical and Electronic Engineering Department, Institute of Advanced Technology, Universiti Putra Malaysia (UPM), Serdang, Malaysia

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_30

Meander line antenna is a type of printed antenna in which the wire structure is embedded on a dielectric substrate to achieve downsizing in size [2]. A meander line antenna is a combination of conventional wire and planer strip line in its most basic form. MLA was developed by a series of studies describes a variety of meander line antenna configurations, including log periodic MLAs, and MLAs with varying vertical segment thicknesses. In this study, we modify the antenna in [4] by using defected ground structure (DGS) [5] instead of coplanar waveguide [7–12] to lower the size of the antenna by half.

2 Antenna Design

2.1 Methodology

The proposed antenna has a small compact dimension of $(34 \times 28 \times 1.1)\ \text{mm}^3$. In addition, instead of coplanar waveguide line, we use the partial ground loaded with defected ground structure (DGS) Fig. 1 shows the proposed antenna. The proposed ground is shown in Fig. 2. A substrate with ($\varepsilon_r = 4.5$) thickness 1.1 mm. The proposed DGS antenna consists of six branches. All the dimensions are labeled in Table 1.

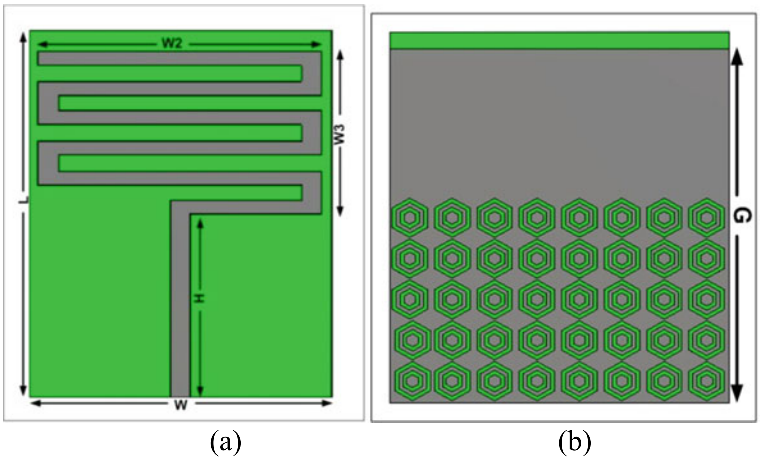


Fig. 1 Geometry of the proposed antenna. **a** Front. **b** Back

Table 1 Proposed antenna (all dimensions in mm)

Parameter	W	L	W2	W3	H	G	a	D	S
Value	28	32	28	15	13	23	3.6	2	0.4

Table 2 Simulations for impedance calculation

H	W	ε	Impedance
1.5	2	4.5	58.4
1	1.5	4.5	54.23
1.1	2	4.5	50

2.2 Impedance Matching

The substrate and feed line thickness have been determined to suit the resistance required 50 impedance calculation software, as shown in Fig. 3, was used to calculate impedance with respect to the height of the substrate and width of the feed line of MLA. Table 2 shows these figures.

2.3 Results and Discussion

Using the −10 dB return loss as a benchmark. The following stages are used to investigate the effect on return loss for various parameter values:

First step: Change the ground height (G) from 17 to 23 mm with step 3 mm increments while keeping the other settings the same. Figure 4 depicts the return loss characteristic. The best return loss S11 value is found when (G) equals 23 mm, 2.5 GHz resonant frequency, and we get bandwidth is 62 MHz. Indicating that the ground height has an impact on the operating frequency and bandwidth. Table 3 shows the various simulation results achieved for altering ground width height.

Second step: Change the feed line's height (H) from (13 to 17 mm) with step 2 mm and without any change in the other parameters. Figure 5 depicts the return loss characteristic. The best value of return loss S11 is determined; when (H) equals 13 mm, the bandwidth is 57 MHz. Table 4 shows the various simulation results achieved with different ground width height.

Figure 6 show the omnidirectional radiation pattern of MLA with gain calculations at main lobe gives 3.21 dB which is very suitable for offered applications.

2.4 SAR Calculations

SAR values are also determined using a10-g reference of human tissue mass [8–10], as shown in Fig. 7. Table 5 illustrates the averaged 10-g SAR when the antenna is relatively close to the body at the aforementioned operating frequency.

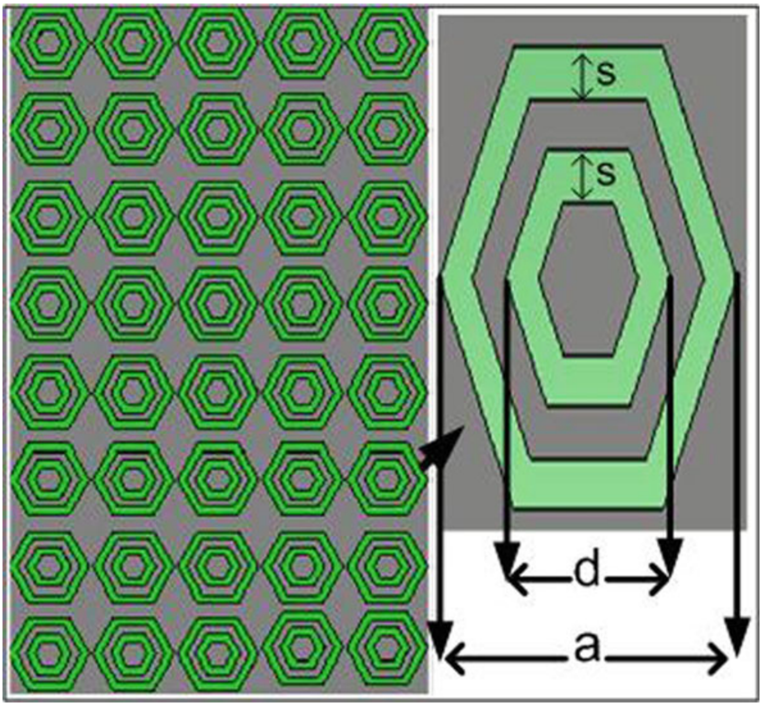


Fig. 2 Geometries of DGS

Table 3 Simulations for different height of the ground (G)

G (mm)	F-start (GHz)	F-stop (GHz)	BW (MHz)
23	2.439	2.501	62
20	2.457	2.517	60
17	2.492	2.551	59

3 Comparison Between Our MLA at 2.5 GHz and MLA at 2.5 GHz in [4]

The authors suggested in the design [4] that 2.5 GHz where HFSS software was used to design and simulate a model antenna as shown in Fig. 8 which consists of a substrate with ($\varepsilon r = 2.5$), and it has thickness equal to 2 mm. A patch antenna with area of 72.625 mm × 72.612 mm × 2 mm. Figure 9 shows the return loss of this antenna. The 3D radiation pattern is shown in Fig. 10.

We see that the previous antenna is a good design for 2.5; our MLA the gain is (3.21 dB). Furthermore, the thickness of our ML is half that of [4]. Table 6 shows the differences between the MLA in [4] and our MLA.

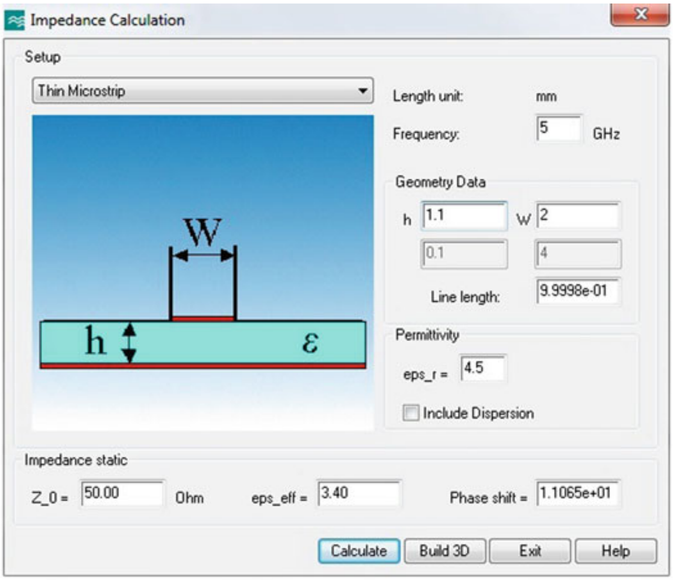


Fig. 3 Impedance calculation

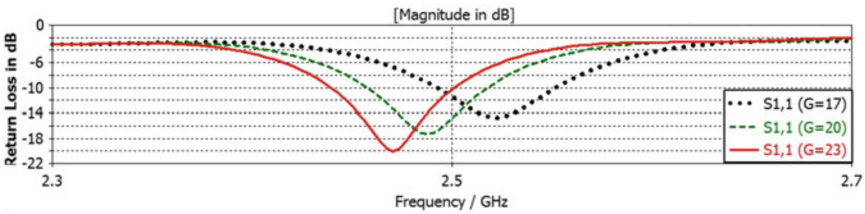


Fig. 4 Return loss (S11) for different ground heights (G)

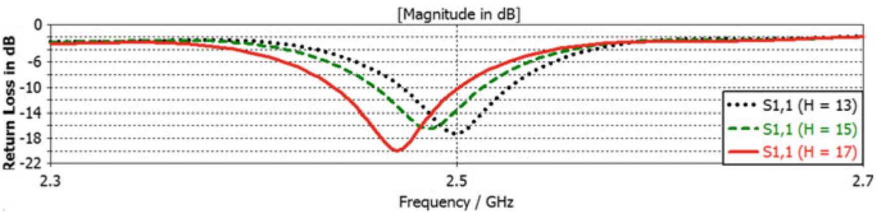


Fig. 5 Effect of different feed line heights (H)

Table 4 Simulations for different height of the feed line (H)

H (mm)	F-start (GHz)	F-stop (GHz)	BW (MHz)
13	2.469	2.526	57
15	2.459	2.512	53
17	2.439	2.501	62

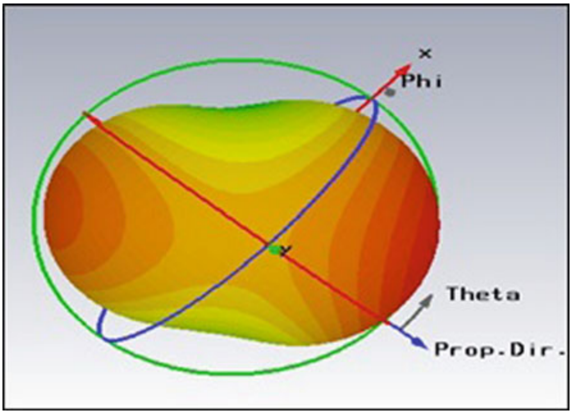


Fig. 6 3D pattern of MLA at 2.5 GHz

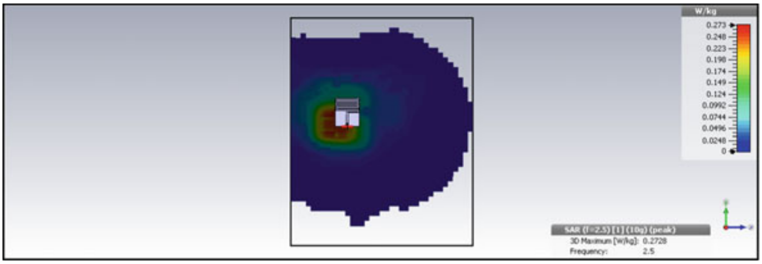


Fig. 7 CTIA-defined talking position

Table 5 SAR values

F	2.5 GHz
SAR (W/kg)—10 g	0.273

4 Conclusion

Meander line antennas have the following advantages: small size, low profile, low cost, and simplicity. Because of these advantages, Meander line antennas are quite common and can be used in a variety of communication systems, including RFID and WLAN. The microstrip patch antenna (meander line antenna MLA) for 2.5 GHz was developed, implemented, and tested in this study using CST Microwave Studio 2014. The diameter of higher for substrate, feed line, and ground is all examined in the MLA parametric studies. We were able to minimize the antenna size by using a defective ground structure. The antenna's SAR values are within acceptable safety limits. Return loss is reduced to −17 dB; bandwidth is increased to 57 MHz, and

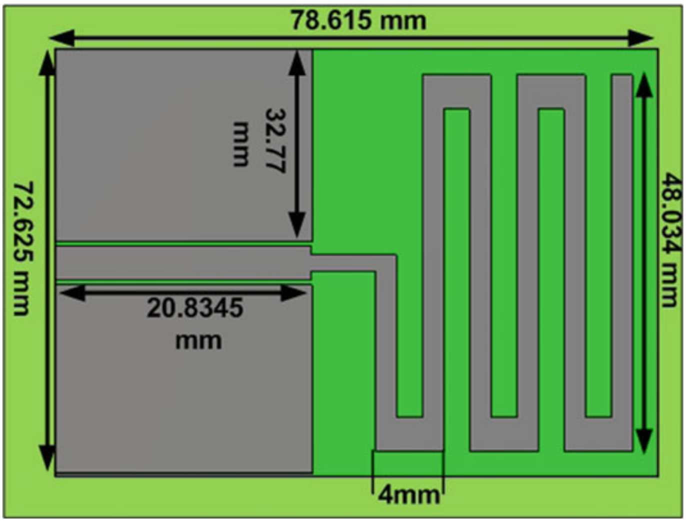


Fig. 8 HFSS design for the MLA in [4]

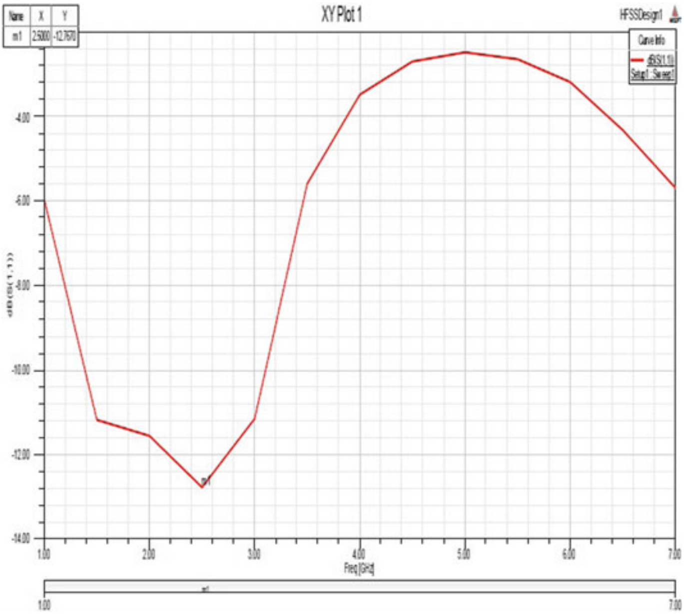


Fig. 9 Return loss for the MLA in [4]

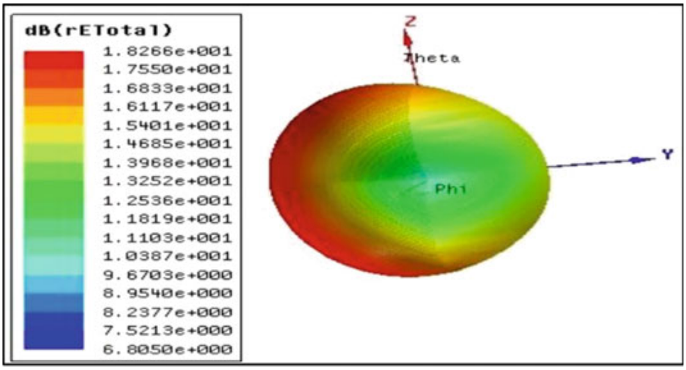


Fig. 10 3D radiation pattern in [4]

Table 6 Difference between the MLA in [4] and our MLA

	Our MLA	MLA in [4]
ε_r	4.5	2.5
Thickness	1.1 mm	2 mm
Return loss S11	−17 dB	−13 dB
No. of line	6	6

gain enhancement is 3.21 dB. The results of a simulation are provided. The proposed antenna has a lot of potential for use in wireless devices like phones and tablets.

References

1. Lin CC, Kuo SW, Chuang HR, 2.4-GHz Printed meander-line antenna for USBWLAN with notebook-PC housing. Student Member, IEEE
2. Frank MC (2002) SkyCross. Inc.
3. Trimukhe MA, Hogade BG (2019) Design and simulation of effects corresponding to reduced conducting area using slotted meandering line for compact ultra-wideband (UWB) antenna. J Commun 14(4):318–323
4. de Oliveira ERS, Chaves Filho ÉF, Sousa MET, de Andrade HD, da Silva Paiva JL (2020) Projeto de uma Antena Dipolo Planar Curvado para Aplicação na Indústria 4.0. Revista Eletrônica de Engenharia Elétrica e Engenharia Mecânica 2(1):103–111
5. Sah S, Tripathy MR, Mittal A (2018) Multiband and miniaturized dual layer antenna incorporated with FSS and DGS. Adv Electromagn 7(1):1–6
6. Bedir Y, Sadiq M, Abdelrazzak M (2015) Design and simulation of meander line antenna for LTE band. Int J Sci Eng Res 6(7)
7. Bedir Yousif MS, Abdelrazzak M (2015) Design and simulation of meander line antenna for LTE communications based on defected ground structure. Ciência e Técnica Vitivinícola 30:15–29
8. Godara LC (ed) (2018) Handbook of antennas in wireless communications. CRC press
9. Azeez HI, Yang HC, Chen WS (2019) Wearable triband E-shaped dipole antenna with low SAR for IoT applications. Electronics 8(6):665

10. Yadav A, Singh VK, Yadav P, Beliya AK, Bhoi AK, Barsocchi P (2020) Design of circularly polarized triple-band wearable textile antenna with safe low SAR for human health. Electronics 9(9):1366
11. Regin R, Obaid AJ, Alenezi A, Arslan F, Gupta AK, Kadhim KH (2021) Node replacement based energy optimization using enhanced salp swarm algorithm (Es2a) in wireless sensor networks. J Eng Sci Technol 16(3):2487–2501
12. Das A, Ghosh A, Sahana S, Singh D, Obaid AJ (2021) An approach to self-reliant smart road using piezoelectric effect and sensor nodes. In: Favorskaya MN, Peng SL, Simic M, Alhadidi B, Pal S (eds) Intelligent computing paradigm and cutting-edge technologies. ICICCT 2020. Learning and analytics in intelligent systems, vol 21. Springer, Cham. https://doi.org/10.1007/978-3-030-65407-8_3

Multi-robot Cooperation and Path Planning Using Modified Cuckoo Search

Bandita Sahu, Pradipta Kumar Das, and Manas Ranjan Kabat

Abstract The paper proposes an innovative approach to solve the cooperation and path planning problem of multiple mobile robots in clutter environment. The main emphasis of the work lies in designing a multi-objective fitness function for stick-carrying robot pairs to compute a collision-free optimal path. The present context of the paper address the multi-robot cooperation and path planning of two pairs of stick-carrying robots that move from a predefined initial location to a pre-assumed goal position by carrying a stick at either end. The basic cuckoo search (CS) algorithm is modified concerning the step size and the scaling parameter at each step of movement of the robot pairs. The modified cuckoo search (MCS) algorithm is implemented with the robot pair to mimic the egg laying behavior of the cuckoo for producing the next generation solution. The proposed algorithm is validated using computer simulation and has been compared with other existing approaches such as ICFA, CS, SDA, and ABCO. Due to its simplicity and efficacy in terms of path optimality, the proposed algorithm produces an optimal solution both in the static and dynamic environment irrespective of the number of obstacles.

Keywords Stick-carrying robot · Multi-robot cooperation · Path planning · Path deviation · Performance

1 Introduction

Cooperation and path planning of multiple mobile robots have become a crucial research problem in the field of robotics [1]. Both the concepts have been paid much attention to solve several social and environmental issues such as object transportation

B. Sahu (✉)
Department of CSE, GIET University, Gunupur, Odisha, India
e-mail: bandita.sahu@giet.edu

P. Kumar Das
Department of IT, VSSUT, Burla, Odisha, India

M. Ranjan Kabat
Department of CSE, VSSUT, Burla, Odisha, India

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_31

in industry, robotics surgery and equipment transfer in healthcare system, detection of planetary movements in space, and disaster management in cluttered unreachable environment [2]. A robot may not be able to do these tasks solely and hence need the help of other robots present in the environment. When more than one robot work together to accomplish a task is called cooperation [3]. For completion of any task, the robots need to move from their position to the desired position. The point at which the transition starts is called as initial position, and the end point of transition is called goal position. Computing the path from the initial to the goal position is called path planning. A path for an autonomous robot can be computed easily. However, planning the path for cooperating robots is complex. Both the cooperation and the planning depend on the environment where the robot resides and has to complete the task. If the environment has motionless obstacle, it is referred to as static environment [4]. But, the environment having moving obstacle is referred to as dynamic environment. Static environment needs less number of computations as compared to the dynamic environment. The cooperation and path planning can be completed using two approaches such as local path planning of multiple mobile robots and global path planning [5]. In local path planning, the robots compute the path in a step-wise manner. It uses the concept of spatial object by considering a local area at each step. However, in global path planning, the robot computes it path as a whole in a known environment. The global approach may not guarantee to produce an optimal solution always [6].

A plethora of algorithms has been proposed to address the issue of path planning and cooperation of the mobile robots. The classical approaches such as potential field method [7], cell decomposition [8], and probabilistic road map [9] method solve the path planning problem of the mobile robots in static environment. But, these are unable to provide an optimal solution in dynamic environment, due to the inefficiency in dealing with the uncertainty [10]. Several measures have been taken for enhancing these classical approaches. Shortest distance algorithm [11] has been proposed to address the path planning problem in static environment without the presence of any obstacle. The authors have considered only an obstacle-free environment to compute the optimal path from the predefined start to the goal state. Hybridization of the existing approaches has also been proposed by using the combined benefits of individual algorithm such as BADE [12], PSO-DV [13], and IPSO-IGSA [14]. All these approaches address the path planning problem only, and the cooperation among the robot is overlooked. Meta-heuristic approaches such as particle swarm optimization (PSO) [15], artificial bee colony optimization (ABCO) [16], and imperialistic competitive firefly algorithm (ICFA) [17] have also been proposed to address the multi-robot path planning and cooperation problem. But they may not produce the optimal solution concerning time and path always as the solution quality depends upon the uncertainty of the environment most of the time. Several nature-inspired techniques such as intelligent water drop algorithm (IWD) [18], cuckoo search algorithm [19], social spider optimization [20], and bat algorithms [21] have been implemented by mimicking the behavior of some natural phenomena to address the path planning problem. However, the objective function of these algorithms is based on the path optimality only and the cooperation part is ignored.

The paper addresses the path planning and cooperation problem of multiple mobile robots in both static and dynamic environment. An optimal and collision-free path is computed using modified CS algorithm. The modification has been realized by tuning the step size and scaling factor dynamically. Two pairs of stick-carrying robot transit from the predefined starting position to a goal position. The main emphasis of the paper is furnished as (i) problem formulation for both path planning and cooperation, (ii) computation of a collision-free path for the robot pairs, (iii) validation of the algorithm by simulating it using programming in C language, and (iv) comparison of the proposed algorithm with other existing approaches.

The rest of the paper is furnished into four sections. Section 2 describes the basic CS algorithm along with the modification done to get the modified CS algorithm. In Sect. 3, the implementation of the proposed algorithm is done with the addressed problem. Section 4 explains the validation of the proposed algorithm through simulation. The conclusion and the future scope of the approach are included in Sect. 5.

2 Cuckoo Search Algorithm

A cuckoo search (CS) algorithm is a nature-inspired evolutionary algorithm inspired by the egg laying behavior of the cuckoo bird. The cuckoo lays their egg in the nest of other host birds. The basic CS algorithm is based on five principles such as (i) each cuckoo lays an egg in a randomly chosen nest, (ii) the egg laid in a host nest represents a solution, (iii) fixed numbers of host nests are assumed, (iv) the host nest having a quality egg is considered as the next generation, and (v) the probability if identifying or destroying the cuckoo's egg or the host bird abandon the best is either 0 or 1. The host nests are identified in a specified area of a fixed radius referred to as egg laying radius (ELR) and expressed as follows:

$$\mathrm{ELR} = E \times \frac{\mathrm{Num}_{\mathrm{laid}}}{\mathrm{Num}_{\mathrm{Total}}}\left(\mathrm{Max}_{\mathrm{egg}} - \mathrm{Min}_{\mathrm{egg}}\right) \tag{1}$$

where E represents the maximum radius, $\mathrm{Num}_{\mathrm{laid}}$ and $\mathrm{Num}_{\mathrm{Total}}$ denote the number of egg laid by a cuckoo and total number of eggs, respectively, and $\mathrm{Max}_{\mathrm{egg}}$ and $\mathrm{Min}_{\mathrm{egg}}$ holds the value of maximum and minimum number eggs than can be laid by a cuckoo. The following equation shows the basic step of computing the solution as a global walk of a current cuckoo $\mathrm{C_i}$ at time $\mathrm{t} + 1$.

$$\mathrm{C_i}(\mathrm{t}+1) = \mathrm{C_i}(\mathrm{t}) + \alpha \oplus \mathrm{levy}(\beta) \tag{2}$$

where the probability distribution of levy flight is

$$\mathrm{levy} \sim u = \mathrm{t}^{\beta} \text{with } 1 < \beta < 3 \tag{3}$$

where α denotes the step size and β is the distribution parameter. The local walk for a cuckoo can also be expressed in the following equation

$$\mathrm{C_i}(t+1) = \mathrm{C_i}(t) + \alpha \oplus \left(\mathrm{C_i^{best}}(\mathrm{t}) - \mathrm{C_i^{best}}(t)\right)\mathrm{levy}(\beta) \tag{4}$$

where $\mathrm{C_i^{best}}(\mathrm{t})$ denotes the best location in the history of the current cuckoo. The equation can further be written using the switching parameter of the host bird from its original nest. The equation for computing the next generation solution of the current cuckoo is represented using a heavy-side function $H(p_c - \varepsilon)$, the switching parameter p_s, and ε as a random value. The local walk is represented as follows.

$$\mathrm{C_i(t+1) = C_i(t)} + \alpha \oplus \mathrm{H}\left(\mathrm{p_s} - \varepsilon\right) \oplus \left(\mathrm{C_i(t) - C_i^{best}(t)}\right) \tag{5}$$

Each time a solution is generated, the fitness value is computed by considering the distance of the host nest from the target position and position of the obstacle present in the environment. It is computed as follows.

$$\mathrm{F(C_i) = comp(C_i, P_{obs})\ \&\&\ D_{ELR}\ \&\&\ D_{i-g}} \tag{6}$$

Such that

$$\mathrm{comp(C_i, P_{obs})} = \begin{cases} 1 & \text{if } \mathrm{C_i == P_{obs}} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

$$\mathrm{D_{ELR}} = \begin{cases} 1 & \text{if } (\mathrm{C_i} - ELR) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

$$\mathrm{D_{i-g}} = \begin{cases} 1 & \begin{array}{l} \text{if } \left(\left|\mathrm{C_i(t) - C_i^{best}(t)}\right| + \left|\mathrm{C_i^{best}(t) - goal}\right|\right) \\ \quad < \left(\left|\mathrm{C_i(t-1) - C_i^{best}(t-1)}\right| + \left|\mathrm{C_i^{best}(t-1) - goal}\right|\right) \end{array} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

The basic CS algorithm suffers from several pitfalls such as (i) low convergence rate, (ii) generation of same solution due to fixed parameters, (iii) repetitive calculation of host nest, and (iv) easy fall into the local optima. To resolve these issues, several measures have been considered with the basic approaches are varying step size, changing the scaling parameter and the penalty function dynamically. The following modifications have been considered in this article for the improvement of basic CS algorithm.

$$\mathrm{p_c = p_c + \frac{rand \times p_c}{n}} \tag{10}$$

$$\alpha = \alpha \times \exp\frac{1}{n} \tag{11}$$

The scaling parameter for computing the probability has been computed using a random function that generates a value between 0 and 1 at each step of calculation. Further, the step size has also been modified to generate new step value in order to explore new host. The modified algorithm can be written as follows.

Algorithm 1: MCS (C)

Input: Initial habitat C

Output: best next generation solution C_i^{best}

Initialize n, n, p_s, α, initial habitat C_i, $Max_{ITER} = 50$ and $t = 1$
Assume $C_i^{best} = C_i$
Evaluate the fitness of each nest $F(C_i)$ using Eq. 6
While ($t <= Max_{ITER}$)
    Select a new nest C_j using Eq. 5
    If $|C_j - C_i| > ELR$
        Drop the selection and made a fresh one
    else
        Evaluate F (C_j)
        If ($F_{profit}(C_i) \leq F_{profit}(C_j)$)
            Set $C_i^{best} = C_j$
    End if
    Update p_c
    Update α
    increase iteration $t = t + 1$
End while

3 Implementation of Proposed Algorithm

The problem addressing the stick-carrying operation is formulated with a pair of robot (R_i,R_j) with the position vector (x_i,y_i) and (x_j,y), respectively, for both the robots. Each time the robot pairs compute a next best position and move with the same velocity and direction. For the cooperation, the velocity, direction of movement, and rotation if any are assumed to be same. At each step, the robot pairs update their position that is assumed as the solution as implemented with the MCS algorithm. The following algorithm describes the computation of the best path from the initial state to the goal state for both the robots.

Algorithm 2: Twin_MCS (C)

Input: Initial habitat C

Output: best next generation solution $C_{i,j}^{best}$

1 Initialize n, p_s, α, initial habitat C_i, $Max_{ITER} = 50$ and $t = 1$
2 Assume $C_{i,j}^{best} = C_{i,j}$
3 Evaluate the fitness of each nest $F(C_i)$ and $F(C_j)$using Eq. 6
4 While ($t <= Max_{ITER}$)
5   Select a new nest C_i' and C_j' using Eq. 5
6   If ($|C_i - C_i'| \geq ELR$ &&$|C_j - C_j'| > ELR$)
7   If ($|C_i' - C_j'| \neq l_{stick}$
8   Drop the selection and made a fresh one
9   else
10   Evaluate F (C_i') andF (C_j')
11   If ($F_{profit}(C_i) \leq F_{profit}(C_i')$ &&)$F_{profit}(C_j) \leq F_{profit}(C_j')$
12   Set $C_{i,j}^{best} = C_{i,j}$
13   End if
14   End if
15   Else
16   Go to step 5 and do a fresh computation
17   End if
18   Update p_c
19   Update α
20   increase iteration $t = t + 1$
21 End while

It is assumed in the algorithm 2 that the robots are separated by a distance same as the stick length l_{stick}. The initial position for the robot pairs are selected from the assumed initial habitat C. The positions are selected as C_i and C_j. Using the equation as mentioned in the previous section, a new next solution is generated for both the robot. If the distance of the new positions is more than the ELR, it is rejected and a new solution is generated. Similarly, if the next position or the solutions generated by the robot pairs do not satisfy the constraint of stick length, the solution is dropped and a fresh calculation is made. The objective function formulated in this algorithm is based on two constraints such as distance constraint between the robot pairs and collision avoidance with the obstacle which is computed in terms of the fitness. It is assumed as a minimization problem that minimizes the path length between the pre-assumed start and goal state.

$$f = w_1 \times f_1 + w_2 \times f_2 \tag{12}$$

Where w_1, w_2 represents the weights and assigned with the value 0.5 and 0.5, f_1 indicates the objective function to compute the distance between the robot pairs and computed as follows.

$$f_1 = \begin{cases} 1 & \text{if} \quad (\left|C_i' - C_j'\right| \neq l_{stick} \\ \left|C_i' - C_j'\right| & \text{otherwise} \end{cases} \tag{13}$$

Similarly, the second objective function describes the fitness of the solution by comparing it with the obstacle position and represented as follows.

$$f_2 = \begin{cases} 1 & \text{if} \quad C_i' \neq C_{obs} \; \mathit{and} \; C_j' \neq C_{obs} \\ D - D_{th} & \text{otherwise} \end{cases} \tag{14}$$

where D_{th} denotes the threshold distance between the obstacle and the robot position, and D denotes the actual distance obtained between the obstacle and the computed position of the robot.

4 Simulation and Result

A simulation environment has been established as a 400 × 550 pixel area on a computer screen. The proposed algorithm along with its competitor has been executed using programming in C language. The environment has been set with 10 obstacles and two pairs of robots. The robots are denoted as circle of radius 5 pixel units. The initial and goal positions are represented with different color circle. The stick is represented as a rectangle of size 10 × 20 pixel area. Experiments have been performed with varying the number of obstacle as 8, 10, 12, and 15. The amount of path traveled by each algorithm like MCS, IVFA, CS, SDA, and ABCO has been computed and shown in Fig. 1. Further, the obstacles are allowed to move for realizing a dynamic environment. To generate a collision-free path concerning the obstacles in motion, all the above-mentioned algorithms are executed and the path traveled in unit pixels is recorded in Fig. 2.

Similarly, the number of turns required reaching the target in the static and dynamic environment is noted in Table 1. The number varies with respect to the number of obstacles. All the competing approaches are executed in the static environment incurred lesser number of turns as compared to the dynamic environment due to the motion of the obstacles. However, irrespective of the number of obstacles and environment, the proposed algorithm produces the optimal path with least number of turns as shown in Figs. 3 and 4. With less number of turns, the proposed algorithm requires less number of steps to reach its target. Figure 5 witnesses the supremacy of the proposed algorithm in terms of incurred steps in static environment. Similarly, the algorithms ICFA, SDA, CS, and ABCO have been executed along with the proposed algorithm in the dynamic environment and require lesser number of steps as shown in Fig. 6. Table 2 illustrates the superiority of the proposed algorithm as compared to other existing approaches in terms of steps. The amount

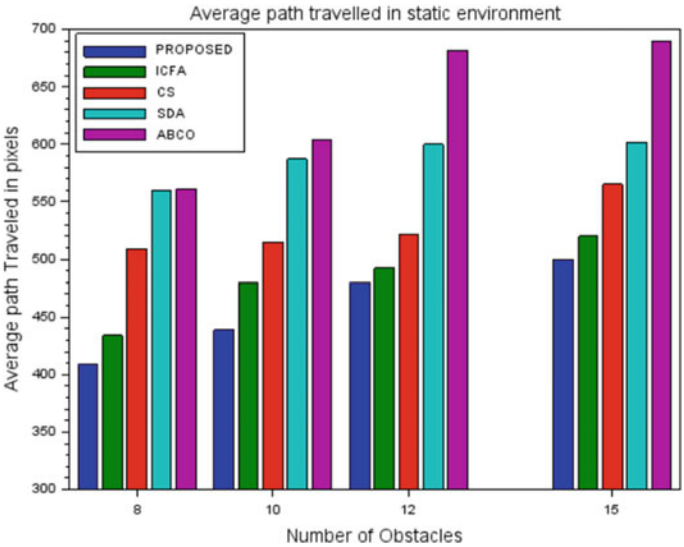


Fig. 1 Average path traveled in static environment

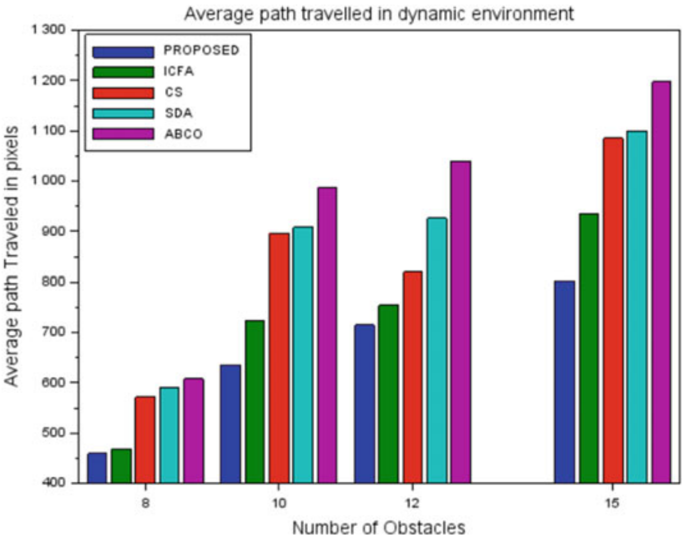


Fig. 2 Average path traveled in dynamic environment

of path deviated by each of the approaches is plotted in Figs. 7 and 8. In both the static and dynamic environment, the amount is minimal for the proposed algorithm.

Table 1 Number of turns require for the robot pairs to reach the target

Algorithms	Environment			
	Static		Dynamic	
	Robot pair1	Robot pair2	Robot pair1	Robot pair2
Proposed	7	11	9	12
ICFA	9	13	13	17
CS	12	16	15	20
SDA	14	19	19	23
ABCO	16	22	20	25

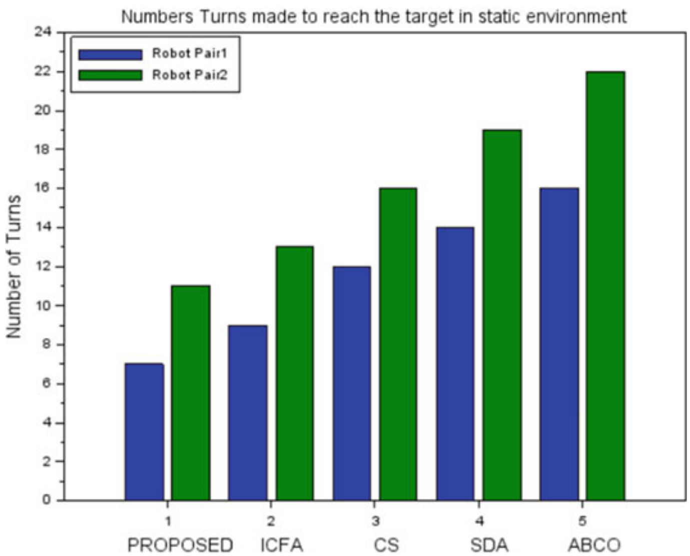


Fig. 3 Turns made by robot pairs in static environment

The overall performance for each of the competing algorithm has been computed by considering the number of steps, turns, and path traveled using the following equation.

$$per = a*steps + b*turn + c*path \tag{16}$$

In the above equation, the variables a, b, and c are assumed as the weighing parameters with the values 3, 2, and 1, respectively. The overall performance computed for

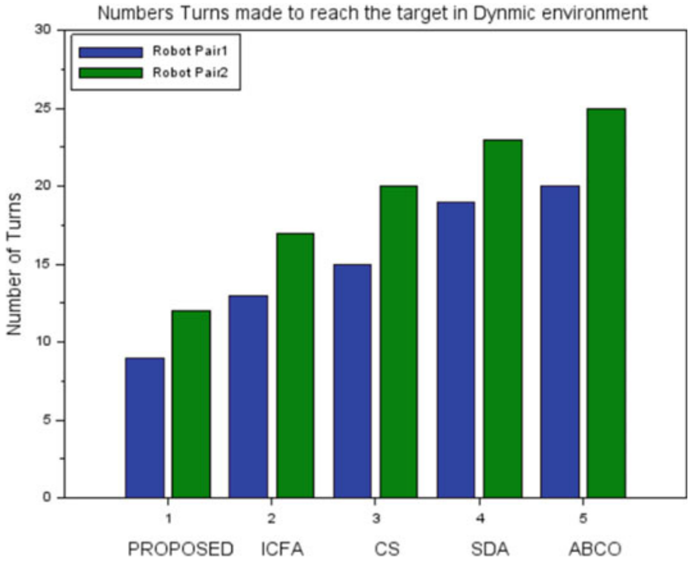


Fig. 4 Turns made by robot pairs in dynamic environment

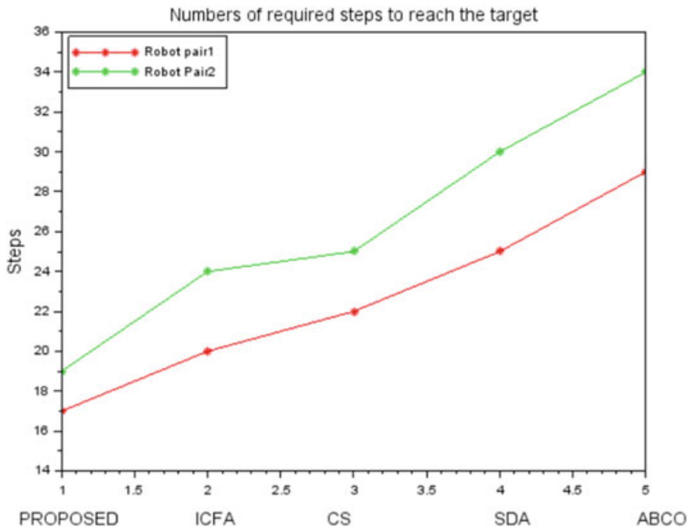


Fig. 5 Steps made by robot pairs in static environment

each of the algorithm is illustrated in Figs. 9 and 10 for the static and dynamic environment, respectively. In both the environment, the proposed algorithm outperforms the competing algorithm.

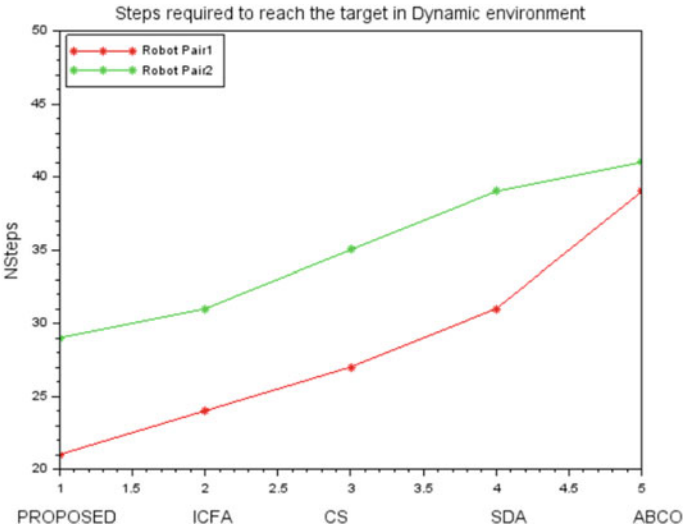


Fig. 6 Steps made by robot pairs in dynamic environment

Table 2 Number of steps require for the robot pairs to reach the target

Algorithms	Environment			
	Static		Dynamic	
	Robot pair1	Robot pair2	Robot pair1	Robot pair2
Proposed	17	30	21	29
ICFA	20	24	24	31
CS	22	25	27	35
SDA	25	30	31	39
ABCO	29	34	39	41

5 Conclusion

The primary focus of the paper is to captivate the problem formulation of multiple mobile robots to transport a stick from one position to another position. The research work is in the framework of nature-inspired techniques to mimic the egg laying behavior of cuckoo. The pitfall of the basic CS algorithm is resolved by tuning the step size and scaling parameter dynamically. The proposed algorithm is validated using computer simulation through programming in C language. The paper justifies the superiority of the proposed algorithm in terms of path optimality, collision avoidance, and performance in both the static and dynamic environment. Further research may be extended for more number of mobile robots to realize the cooperation among them as the proposed paper is limited to the synchrony of two robots only.

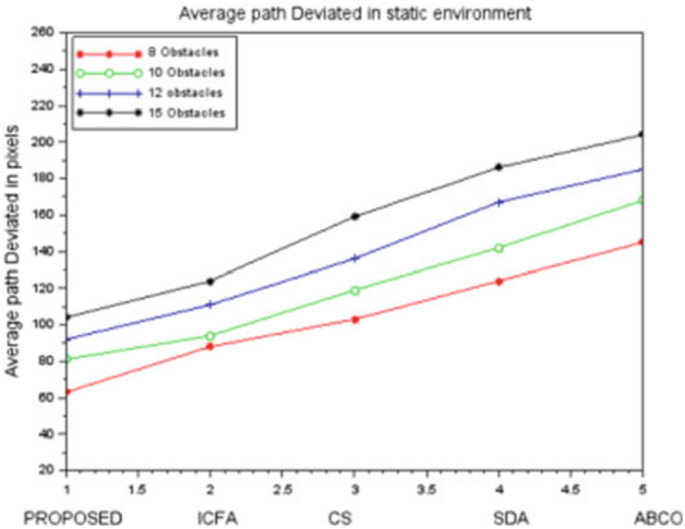


Fig. 7 Path deviated in static environment

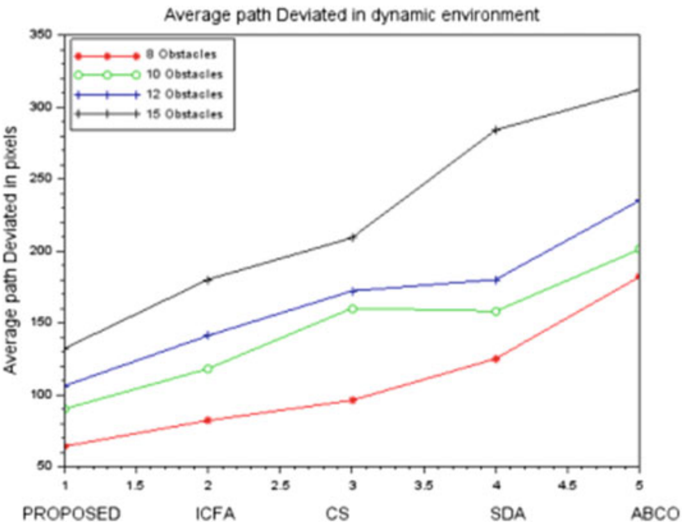


Fig. 8 Path deviated in dynamic environment

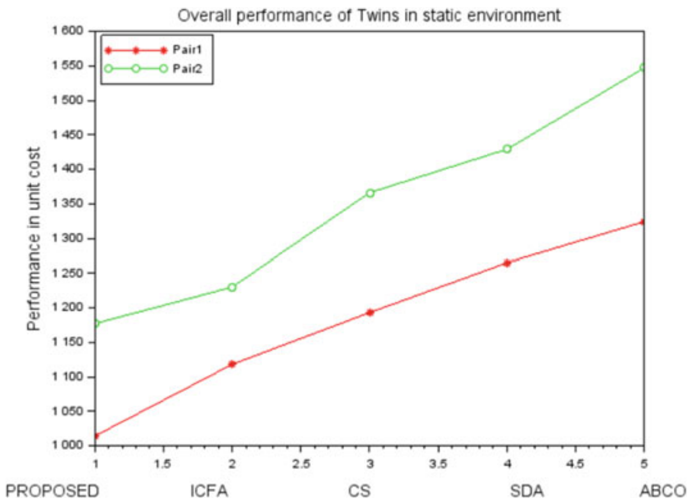


Fig. 9 Overall performance in static environment

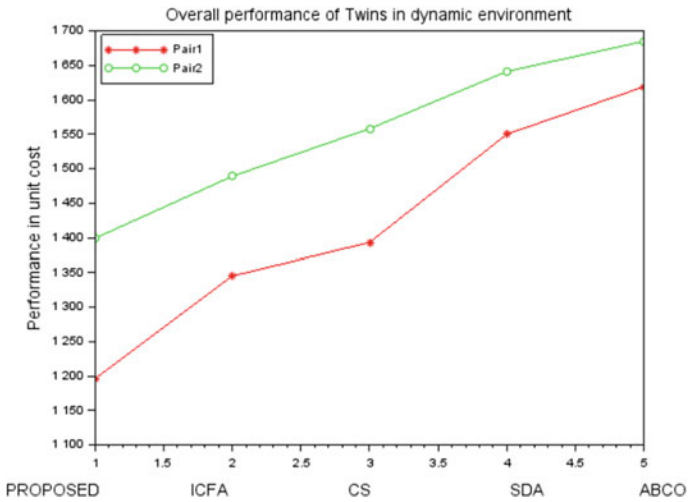


Fig. 10 Overall performance in dynamic environment

References

1. Olcay E, Schuhmann F, Lohmann B (2020) Collective navigation of a multi-robot system in an unknown environment. Robot Auton Syst 132:103604
2. Rajendran P et al (2021) Strategies for speeding up manipulator path planning to find high quality paths in cluttered environments. J Comput Inf Sci Eng 21(1):011009
3. Hamed O, Hamlich M (2020) Improvised multi-robot cooperation strategy for hunting a dynamic target. EAI Endorsed Trans Internet Things 6(24):e5

4. Chandrashekhar A, Himam Saheb S, Pavan Kishore ML (2021) Investigation of the static and dynamic path planning of mobile and aerial robots. Comput Networks Inventive Commun Technol, 1033–1044
5. Ab Wahab MN, Nefti-Meziani S, Atyabi A (2020) A comparative review on mobile robot path planning: classical or meta-heuristic methods? Annu Rev Control
6. Zhang H-y, Lin W-M, Chen A-X (2018) Path planning for the mobile robot: a review. Symmetry 10(10):450
7. Vadakkepat P, Tan KC, Ming-Liang W (2000) Evolutionary artificial potential fields and their application in real time robot path planning. In: 2000 Proceedings of the congress on evolutionary computation. CEC00 (Cat. No. 00TH8512). vol 1, IEEE
8. .Kloetzer M, Mahulea C, Gonzalez R (2015) Optimizing cell decomposition path planning for mobile robots using different metrics. In: 2015 19th international conference on system theory, control and computing (ICSTCC), IEEE
9. Santiago RM et al (2017) Path planning for mobile robots using genetic algorithm and probabilistic roadmap. In: 2017 IEEE 9th international conference on humanoid, nanotechnology, information technology, communication and control, environment and management (HNICEM). IEEE, 2017.
10. Liu Z et al (2020) Prediction, planning, and coordination of thousand-warehousing-robot networks with motion and communication uncertainties. IEEE Trans Autom Sci Eng
11. Ali AA et al (2016) An algorithm for multi-robot collision-free navigation based on shortest distance. Rob Auton Syst 75:119–128
12. Meng X, Gao X, Liu Y (2015) A novel hybrid bat algorithm with differential evolution strategy for constrained optimization. Int J Hybrid Inf Technol 8(1):383–396
13. Das PK et al (2016) A hybrid improved PSO-DV algorithm for multi-robot path planning in a clutter environment. Neurocomputing 207:735–753
14. Das PK, Behera HS, Panigrahi BK (2016) A hybridization of an improved particle swarm optimization and gravitational search algorithm for multi-robot path planning. Swarm Evol Comput 28:14–28
15. Song B, Wang Z, Zou L (2021) An improved PSO algorithm for smooth path planning of mobile robots using continuous high-degree Bezier curve. Appl Soft Comput 100:106960
16. Das P et al (2015) Arduino based multi-robot stick carrying by artificial bee colony optimization algorithm. In: 2015 proceedings of the third international conference on computer, communication, control and information technology (C3IT), IEEE
17. Sadhu AK, Rakshit P, Konar A (2016) A modified imperialist competitive algorithm for multi-robot stick-carrying application. Robot Auton Syst 76:15–35
18. Salmanpour S, Motameni H (2014) Optimal path planning for mobile robot using intelligent water drops algorithm. J Intell Fuzzy Syst 27(3):1519–1531
19. Saraswathi M, Murali GB, Deepak BB (2018) Optimal path planning of mobile robot using hybrid cuckoo search-bat algorithm. Procedia Comput Sci 133:510–517
20. Kavitha S, Venkumar P (2020) A vibrant crossbreed social spider optimization with genetic algorithm tactic for flexible job shop scheduling problem. Meas Control 53(1–2):93–103
21. Guo J, Gao Y, Cui G (2015) The path planning for mobile robot based on bat algorithm. Int J Autom Control 9(1):50–60

Technology Framework for Building Educational Augmented Reality Applications

Hung Ho-Dac, Van Len Vo, and Tuan Anh Tran

Abstract Applying technology to education to improve efficiency is one of the inevitable trends. In higher education, augmented reality (AR) is strongly used due to its advantages. AR provides the ability to apply simulation models in teaching to increase the visualization as well as extend conditions for students to practice more and more. Derived from that fact, we have research motivation for applying AR technology in education, especially for higher education. In this work, we propose a procedure to build up an educational AR application as well as a related technology framework. We also build up an AR application following our proposed procedure and technology framework. This application allows users to interact with machine elements. Furthermore, we model machine elements in automobile major which are available on GitHub and free.

Keywords AR · Simulation · Technology framework

1 Introduction

Nowadays, augmented reality (AR)—one of the key areas of Industrial Revolution 4.0 is developing and applying strongly in many different fields. The deployment of 5G networks combined with cloud technology makes the deployment of AR applications into social life easier than ever. In teaching, AR is strongly used to improve the quality of education and training because of its advantages, especially the ability to apply simulation models in teaching to increase the visualization. Applying AR to education provides students with a new and more effective learning platform than traditional teaching methods where students can practice through interaction to gain the ability

H. Ho-Dac (✉) · V. Len Vo · T. Anh Tran
Thu Dau Mot University, Binh Duong Province, Thu Dau Mot, Vietnam
e-mail: hunghd@tdmu.edu.vn

V. Len Vo
e-mail: lenvv@tdmu.edu.vn

T. Anh Tran
e-mail: 1825202010102@student.tdmu.edu.vn

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_32

to experience, observe, and innovate through full practical exercises with the support of AR technology. From there, it helps students improve their professional skills and educational efficiency.

Currently, most students own at least one smartphone. So the combination of AR, smartphone, and education becomes easy and convenient. Augmented reality offers a combination of real, user-viewed, and computer-generated virtual scenes. This is the augmentation of the real world by adding in real space a place, space, thing, or event. This new approach enhances the effectiveness and attractiveness of teaching and learning [1].

Due to that fact, we realize that building procedures and technology frameworks for AR application in education is necessary. In this work, our contributions are as follows:

(i) We propose a procedure and technology framework for building up AR applications in education.
(ii) We build up an open-source AR application following our proposed procedure and technology framework which allows users to interact with machine elements.
(iii) We model machine elements in automobile major and publish them free.

Our source code and model can be easily found at https://github.com/lamtacta2/NCKH2020-2021. The rest of the paper is organized as follows: Sect. 2 surveys the related work. Section 3 introduces the procedure and technology framework. We present our AR application in Sect. 4. We conclude our work with future work in Sect. 5.

2 Related Work

There are many studies in applying AR to education. In [2], Kesim and Ozarslanb introduce the application of AR in education as well as the possibilities of AR. They discuss some key technologies in AR and their impacts. Wu et al. [3] conduct a survey and discuss technological, pedagogical, and learning issues related to the implementation of AR in education.

In another work, Chen et al. [4] survey studies published between 2011 and 2016 and provide state of art on AR in education. They also discuss trends and opportunities of AR in education. Lee [5] explains in his literature review how AR can be applied to education and the potential impact on future education. In [6], Wang et al. present the development of AR in education. They also discuss the potential of AR tools in teaching and learning.

VR and AR are increasingly being used in the medical field [7, 8] as VR surgical simulation systems that provide trainee or inexperienced surgeons with surgical training when operating a VR surgical simulation system, and you can simulate the realism of the actual operation and reduce the incidence of errors during the actual operation in the future. There has been a steady growth in the use of virtual reality

(VR) in health care [9]. In the teaching of human anatomy [10], three-dimensional stereoscopic visual effects can be presented through VR to understand the relative positional relationship between objects. For clinicians or medical students, in the understanding of human structure and learning has very good potential application. In healthcare education, intravenous injection can also be learned through VR and AR [11].

3 Procedure and Technology Framework

There are many architectures for AR systems [12, 13]. From our perspective and best knowledge, we propose a skeleton for AR system on mobile phones including but not limited to elements (Fig. 1): mobile phones, user interfaces, peripherals, data/models, and applications.

We also propose the overall procedure for building up AR application in education as in Fig. 2. We will explain in detail our procedure and technology framework in this section. We propose Vuforia [14], SOLIDWORKS [15], Blender [16], Unity [17], Visual Studio [18], and PowerPoint [19] for technology framework. Vuforia is developed by PTC company. This technology helps to connect the phone's camera to the software. Retrieve data from camera input and process. Currently, this technology supports 3D scanning on iOS, tracking images, tracking plants, etc. We used it to deal with the function of connecting the phone's camera to the software and Tracking Image. SOLIDWORKS is software that supports mechanical design, construction, and simulation. This software is quite easy to use and convenient. We used it to create small elements and combine them to form large elements. To integrate big elements into Unity as 3D models, there are two solutions: (i) use a support tool but will pay fees like T-Link, (ii) use free software to convert the format of the model to a

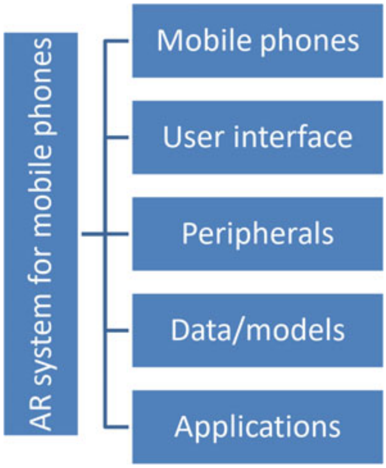


Fig. 1 Skeleton for AR system on mobile phones

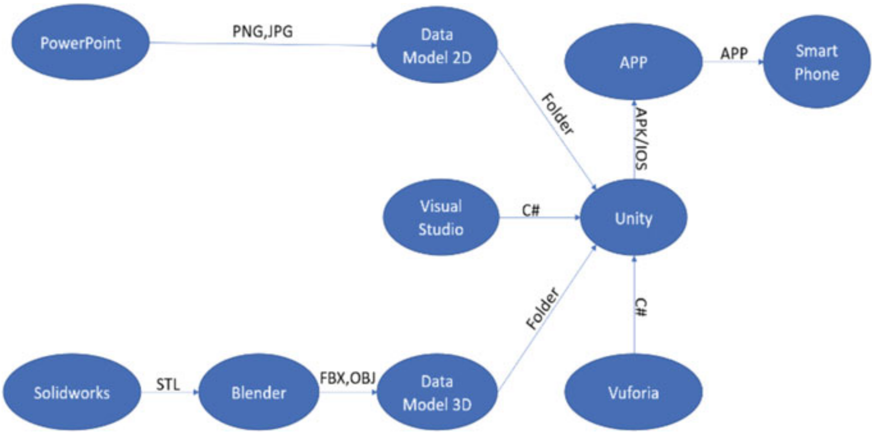


Fig. 2 Overall procedure and technology framework for building educational AR application

suitable format for Unity. Blender is a cross-platform tool. It also has relatively small memory and drives requirements compared to other 3D creation suites. Blender can be used to create 3D visualizations such as still images, 3D animations, VFX shots, and video editing. We chose Blender because it is free and easy to use. Unity is the software; we choose to connect things together and supports exporting files to many different formats such as iOS, Android, etc. To connect and export files to iOS or Android, there are many supported software applications but now most prominent are Unreal Engine and Unity. We chose Unity because it uses C# which is easy to learn and use. Visual Studio is the official IDE of Microsoft. We use this IDE to develop scripts for Unity, This IDE is quite easy to use in terms of ease of learning and has long-term support. PowerPoint 2019 is a presentation aid software but can also be used to design 2D because it is quite easy to use and has quite a full range of features for 2D design and it also helps to export files as PNG and JPG. So we chose this software to generate 2D data for the project.

To build a 2D model, we can use it in two ways: (i) drawing by hand, (ii) using supporting software. With the first way, drawing by hand will require drawing skills and take quite a lot of time while the second method is simple. So here we choose the second option. Currently, there are many supporting software applications such as Paint, Adobe Photoshop, etc. But due to the economic and educational purpose, we use PowerPoint. PowerPoint software has the feature to help users export files as PNG or JPG and the software is quite easy to use and familiar to us. To create a 2D model, we follow these steps:

- Step 1: Take a picture of the 3D model.
- Step 2: Create a slide in PowerPoint.
- Step 3: Design the frame for the 2D model.
- Step 4: Insert the photo captured in step 1 into the frame just designed in step 3.

- Step 5: Export the slide as PNG or JPG and save it to a file to create data for the system.
- Step 6: Print the photo just created in step 5 to create a card.

To build a 3D model, we use SOLIDWORKS software to draw the elements and then assemble them to form a complete model. Doing each detail separately will help the system to have large data and more flexibility. Once we have drawn the details of a large model, we start to put them together. To draw the elements, we use SOLIDWORKS part and assembler we use Assembly of SOLIDWORKS. After the assembly is complete, we will export the file as STL, we start to put the newly created STL files into Blender by the Import command in Blender. After importing into Blender we export the file as FBX or OBJ. FBX will form the discrete parts and OBJ will combine those discrete parts into one block. That means the FBX file will have but the OBJ file is inside and the OBJ file only has it. Once we have the FBX or OBJ file, we bring Unity to drag objects from outside into Unity or copy or move them to the folder where the project is stored. When brought into Unity, 3D models will lose colors. To color the 3D model, we can color in Blender or Unity. In Blender, we have two ways to use Texture Image or mesh color. If using a Texture Image, an image with a color matching each point on the 3D model is needed. This is quite difficult and time consuming. If you color in a grid, it is very time consuming and detailed models will be difficult to paint and causes color smudges easily. So we chose to fill Unity with the FBX file format.

We use Unity to connect all things together in an application. In addition, we integrate Vuforia technology and peripherals control into the application. The detail source code and models can be found easily at our repository on GitHub (https://github.com/lamtacta2/NCKH2020-2021).

4 Application

We build up an AR application following our proposed procedure and technology framework which allows users to interact with machine elements. When the camera of a mobile phone captures a card representing the 2D model of a specific machine element, the corresponding 3D model will be loaded. Users can interact with the 3D model in many ways including but not limited to: rotate, zooming, moving, etc. Our application also provides basic information for each machine element.

We do believe that with the proposed procedure and technology framework, the time and effort required to build an AR application for education will not be too much. We complete our application in about 1 month with 10 3D models and ten corresponding 2D models (Figs. 3, 4 and 5).

Fig. 3 Educational AR application in automobile major screenshot#1

Fig. 4 Educational AR application in automobile major screenshot#2

Fig. 5 Educational AR application in automobile major screenshot#3

5 Conclusion

In this work, we propose a procedure to build up an educational AR application as well as a related technology framework. We also build up an AR application following our proposed procedure and technology framework. This application allows users to interact with machine elements. Furthermore, we model machine elements in automobile major which are available on GitHub and totally free. In our further work, we will apply more complex interactions and more models to our repository.

Acknowledgements This research is funded by Thu Dau Mot University under grant number DT.21.1- 076. Also, we would like to thank Institute of Engineering Technology (TDMU) for the support of time and facilities for this study.

References

1. Earle RS (2002) The integration of instructional technology into public education: promises and challenges. Educ Technol 42(1):5–13
2. Kesim M, Ozarslan Y (2012) Augmented reality in education: current technologies and the potential for education. Procedia Soc Behav Sci 47:297–302
3. Wu HK et al (2013) Current status, opportunities and challenges of augmented reality in education. Comput Educ 62:41–49
4. Chen P et al. A review of using augmented reality in education from 2011 to 2016. Innovations in smart learning, pp 13–18
5. Lee K (2012) Augmented reality in education and training. TechTrends 56(2):13–21
6. Wang M et al (2018) (2018) Augmented reality in education and training: pedagogical approaches and illustrative case studies. J Ambient Intell Humanized Comput 9(5):1391–1402
7. Dascal J et al (2017) Virtual reality and medical inpatients: a systematic review of randomized, controlled trials. Innov Clin Neurosci 14(1–2):14
8. Van Krevelen DW, Poelman R (2010) A survey of augmented reality technologies, applications and limitations. Int J Virtual Reality 9(2):1–20
9. Moline J (1997) Virtual reality for health care: a survey. Virtual reality neuro-psycho-physiology, pp 3–34
10. Lee S et al (2013) Augmented reality intravenous injection simulator based 3D medical imaging for veterinary medicine. Vet J 196(2):197–202
11. Barsom EZ, Graafland M, Schijven MP (2016) Systematic review on the effectiveness of augmented reality applications in medical training. Surg Endosc 30(10):4174–4183
12. Chatzopoulos D et al (2017) Mobile augmented reality survey: from where we are to where we go. IEEE Access 5:6917–6950
13. Mota JM et al (2018) Augmented reality mobile app development for all. Comput Electr Eng 65:250–260
14. "Vuforia Engine | Create AR Apps And AR Experiences | PTC". Ptc.Com, 2021, https://www.ptc.com/en/products/vuforia/vuforia-engine.
15. SOLIDWORKS (2021) https://www.solidworks.com/. Accessed 01 Apr 2021
16. Foundation Bl (2021) Blender.Org—Home of the blender project—Free and open 3D creation software. Blender.Org. https://www.blender.org/
17. Technologies Unity (2021) Unity—Unity. Unity. https://unity.com/

18. Visual Studio IDE, Code Editor, Azure Devops, App Center—Visual Studio. Visual Studio, 2021. https://visualstudio.microsoft.com/
19. Microsoft powerpoint slide presentation software | Microsoft 365. Microsoft.Com, 2021. https://www.microsoft.com/en/microsoft-365/powerpoint.

Multiband Handheld Antenna with E-shaped Monopole Feeding

Mohammed Sadiq, Nasri Bin Sulaiman, Maryam Biti Mohd, and Mohd Nizar Hamidon

Abstract An antenna for wireless communication applications, which is a simple multi-band planar antenna is offered. This work introduces an open slot antenna supplied by an E-shaped monopole for use with mobile and wireless LAN services. The objective of this paper is to design printed antennas suitable for use in LTE mobile stations. The big challenge is to obtain small frequencies from an antenna of a small size because the inverse relationship between antenna size and frequency, safe for human use, suitable for use in DVB, operates for most of the mobile applications, and with a wide bandwidth. The antenna size is $42 \times 33.8 \times 1.5$ mm^3. We added three branch lines for the proposed antenna to accommodate the Digital Video Broadcasting DVB. The antenna's length is adjusted to 500 MHz to support DVB bands; it runs in five bands: 470–530 MHz, 666–750 MHz, 862–980 MHz, 1.37–2.88 GHz, and 3.15–3.52 GHz. The SAR computations are performed using the commercial program CST 2014. It is worth noting that the experimental measurements were compared to the simulation results and show a high degree of compatibility.

Keywords MLA · DVB-H · SAR

M. Sadiq (✉)
Dhi-Qar Investment Commission, Nasiriyah, Iraq
e-mail: engmohammedalkaabi@gmail.com

N. B. Sulaiman
Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, 43400 Putra, Malaysia
e-mail: nasri_sulaiman@upm.edu.my

M. B. Mohd
Department of Electrical and Electronic Engineering, Universiti Putra, Putra, Malaysia
e-mail: maryam@upm.edu.my

M. N. Hamidon
Electrical and Electronic Engineering Department and Deputy Director of Institute of Advanced Technology at University Putra Malaysia (UPM), Putra, Malaysia
e-mail: mnh@upm.edu.my

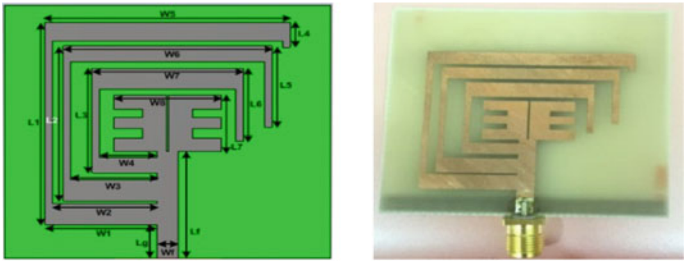
R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_33

1 Introduction

Different requirements for mobile phone antenna design have been anticipated and expanded in response to the support and enhancement of multimedia functionality by mobile communication systems. Numerous mobile users are interested in receiving L-band broadcasts via DVB-H (Digital Video Broadcasting-Handheld) and T-DMB (Terrestrial Digital Multimedia Broadcasting). LTE, the 4th (Fourth Generation of Cellular Networks), is also meant to provide multimedia services to any location, at any time. The LTE standard is expected to operate in bandwidths between 1.4 and 20 MHz, in a range of frequency bands between (400 MHz and 4 GHz) [1–3]. Numerous studies have been conducted in order to establish an antenna architecture that fits the DVB-H antenna requirements for handheld devices [1, 2]. Although ideas utilizing control circuits for miniaturization were developed [4–7], but there are some problems such as the reduced antenna efficacy, and increased prices. Magneto-dielectric materials have been primarily used to reduce the size of antennas [8–10], but they have a number of disadvantages, including high loss, low radiation efficiency, and a high cost when compared to conventional dielectric materials [11–15]. Seong et al. [12] offer a small antenna that incorporates a U-shaped slot printed in rectangular monopole for DVB integration. To cover DVB-H, T-DMB, and GSM 900, the antenna takes up only 50×27 mm^2. Langley et al. [13] presented a compressed multi-band tunable system with a vector as an effective tuning component to contain a wide of mobile bands UMTS, GSM (85,090,018,001,900) by regulating the capacitance between two broadcast bands and the gap in the slotted PIFA, radio frequency FM 76–108 MHz, and mobile television DVB-H 470–702 MHz with size of $40 \times 15 \times 8$ mm^3. Yang et al. [14] designed a planar meander monopole antenna. From 440 to 1350 MHz, wideband impedance characteristics are achieved utilizing parasitic strips and a sleeve feed. It is ideal for LTE, DVB-H, and GSM (850,900) applications in mobile phones measuring 40×100 mm^2. This article describes an internal antenna that consists of a monopole have 3-branch that can span a wide frequency spectrum with DVB and LTE. The return loss for the suggested antenna is a -10 dB, and bandwidth of 470–533 MHz, 650–750 MHz, 869–993 MHz, 1.372–2.884 GHz, and 3.142–3.525 GHz, which means it covers UMTS (2100), ISM (2450), PCS (1900), DCS (1800), and GSM (900), the majority bands of LTE with a diameter of $42 \times 33.8 \times 1.5$ mm^3.

2 Antenna Design

The proposed antenna with compact dimensions of $(42 \times 33.8 \times 1.5)$ mm^3 is a planar printed antenna and can joint into a small and smart mobile device. The proposed antenna design in this study is illustrated in Fig. 1. The substrate of antenna is constructed from FR4 (4.4), thickness 1.6 mm, and the magnitude of loss-tangent equal to (0.025), as illustrated in Fig. 1.



(a) (b)

Fig. 1 The geometry of propose antenna

Our presented patch antenna is consisting a two E-shape planar back-to-back monopole and three sub-lines. At 900 MHz, the first sub-line work as a monopole radiator as well as a bore resonator supplied from E-shape monopole resonator supplied from E-shape monopole antenna that emits in the higher frequency bands. The second and third sublines extend the route of the surface current, so lowering the resonance frequency. The first line's electrical length of 61.25 mm is set to emit in at 900 MHz (869–993 MHz). The second and third lines' electrical lengths of 83.85 mm and 89.8 mm, respectively, are tuned to resonate at 700 MHz (650–750 MHz). To further understand the implications of various dimensions parameters, a complete parametric analysis was conducted. Setting the value of some parameters for the resonant frequency can be done in a step-by-step method [16]. The first step is to design the antenna's dimensions, as illustrated in Fig. 1, with the beginning value of 0.5 mm for all of the antenna's dimensions. The settings are specified as variables to demonstrate how they affect the MLA's bandwidth and gain.

CST Microwave Studio 2014 was used to simulate the suggested antenna. The comparison in terms of the actual and modeled data for the return loss (S11) was illustrated in Fig. 2. According to simulation and testing data, the antenna covers many of the foregoing mobile and wireless application bands. The antenna operates in three bands (674–750 MHz, 869–993 MHz, and 674–750 MHz) with a return loss reference of −10 dB (1.4–3.13 GHz). The length of both the second and first lines is adjusted to control the first and second resonant frequencies, respectively, while the size of monopole radiators and branch lines is adjusted to regulate the third resonant frequency.

3 SAR Calculations

As mobile phone usage increases, so does research into the potential health dangers posed by electromagnetic (EM) fields generated by wireless terminals. When a cell phone is held close to the head or hand, a multitude of factors can affect the EM interaction. SAR, or specific absorption rate, is a widely used metric for assessing

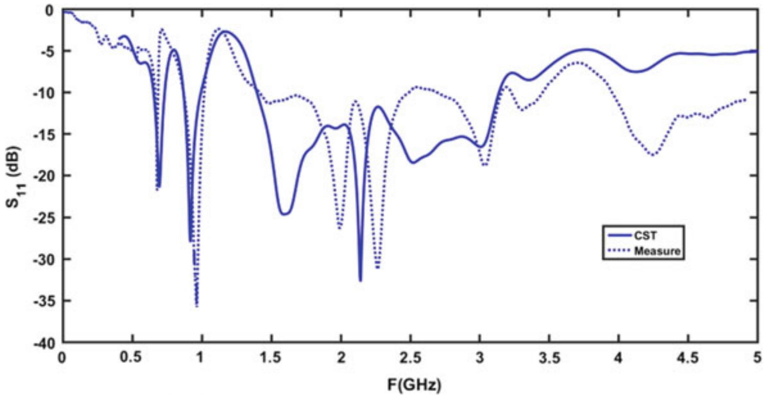


Fig. 2 The proposed antenna's return loss was simulated and measured

exposure. As a result, numerous rules and recommendations have been enacted to regulate mobile handset radiation exposure, with the goal of reducing SAR while also increasing antenna system efficiency.

IEEE C95.1: 2005 has been amended to increase the SAR limit to 2 W/kg across any 10 g of tissue [17], equivalent to the threshold recommended in ICNIRP guidelines [18]. The cellular phone model's output power must be specified before SAR can be simulated. The cellular phone's output power is set to 500 mW in this study, with frequencies of operation of 0.7, 0.9, 1.8, and 2.1 GHz. The model of Hugo Voxel (structure of antenna near the human head model) is depicted in Fig. 3. The SAR values are derived using a 10 g human tissue mass as a reference. The SAR computations and the Hugo Voxel model are performed using the commercial program CST 2014 [19], and the relative permittivity and conductivity of the tissues involved as it is specified in [20–23]. Table 1 explains the 10 g-SAR values at the previously mentioned frequencies band when the antenna is relatively close to the human body.

The antenna complies with standards of ICNIRP and IEEE C95.1: 2005. The Specific Absorption Ratio values are modified by varying the space between the antenna and the head due to the scalability of the results when the power level fluctuates.

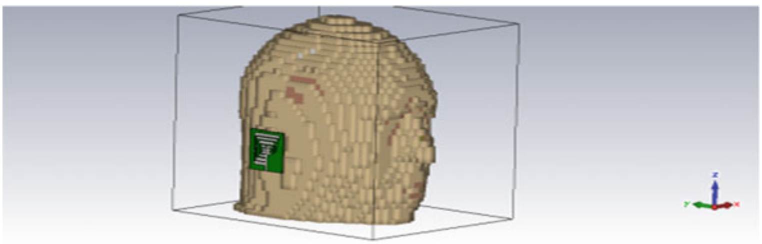


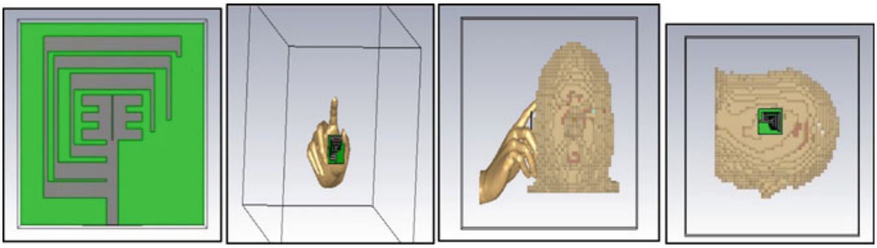
Fig. 3 Structure of the antenna modeled after a human head (Hugo voxel model)

Table 1 Values of SAR, and the human model impacts of on antenna characteristics

F in GHz	Human	Without human	Human (dB)	Without human (dB)	SAR (W/kg) (10g)
	Efficiency %		S11		
0.7	77	80	−14	−20	0.59
0.9	82	92	−25	−27	0.41
1.8	84	86	−17	−16	0.72
2.1	79	80	−14	−33.9	0.92

As illustrated in Fig. 4, Free space, Hand Mode (Browsing), speaking position with hand near from head (the space between human head and antenna 0.5 mm), and speaking position were proposed by the Cellular Telecommunications and Internet Association (CTIA) as four different test sites for the body and mobile phone [18].

Figure 5 illustrates the return loss for four different scenarios of the antenna. The hand and head have the primary impact of causing minimal impedance matching shift and degradation. For real-world mobile phone applications, however, impedance matching across the operating bands is still necessary.



(1) (2) (3) (4)

Fig. 4 CTIA-defined. **1** Free space, **2** hand mode (browsing), **3** speaking position with hand, and **4** speaking position

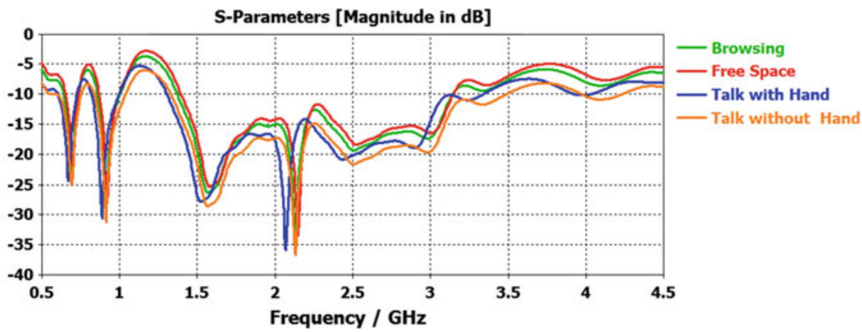
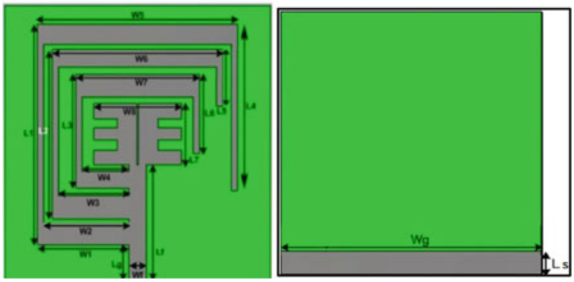


Fig. 5 S11 at the four CTIA positions



(1) Front view (2) Back view

Fig. 6 DVB antenna

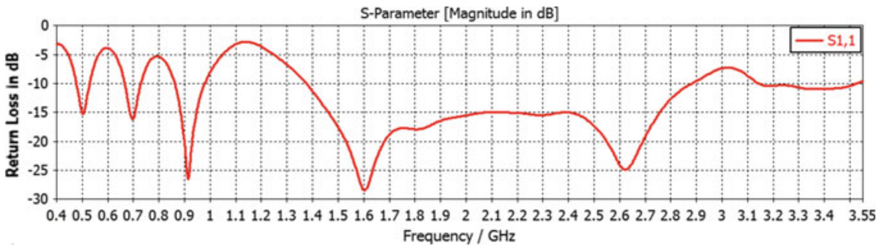


Fig. 7 The suggested DVB antenna's return loss was simulated and measured

4 Antenna Design with DVB Band

To accommodate the Digital Video Broadcasting (DVB band), the suggested patch antenna is changed by adding three sub-lines of varying lengths, as illustrated in Fig. 6. The third sub-line enlarges the channel through via which the surface current circulates, hence decreasing the resonance frequency. The dimensions are increased to accommodate DVB bands, and the antenna measure are increased to $(42 \times 33.8 \times 1.5)\ \text{mm}^3$.

The return loss simulation results are shown in Fig. 7. The antenna's coverage of all preceding bands, as well as the DVB band, is ensured by the simulation results. The antenna works in the following bands: 470–530 MHz, 666–750 MHz, 862–980 MHz, 1.37–2.9 GHz, and 3.15–3.52 GHz, utilizing a −13 dB reference.

5 Conclusion

The invention of a compact-planar antenna that is suitable for the majority of currently available mobile, wireless communication services, DVB, and ISM service is described. The Specific Absorption Ratio values are within acceptable safety standards. The effect of the human body at antenna efficiency was also taken into account.

The antenna is smaller than other antennas that have been published. The CST simulator was used to simulate the antenna, and the photolithographic process was used to construct it. The simulated and experimental results show a high level of agreement.

References

1. Page TA, Jahagirdar TZ, Mujawar TU (2018) DVB-SH-A transmission in Gaussian and mobile fading channels. Int J Eng Technol 3(1):1–12
2. Singh AK, Tripathy MR (2018) M shaped ultra wide band monopole antenna for TVWS CPE. In: 2018 5th international conference on signal processing and integrated networks (SPIN) (pp 240–243), IEEE
3. Plonis D, Katkevičius A, Gurskas A, Urbanavičius V, Maskeliūnas R, Damaševičius R (2020) Prediction of meander delay system parameters for internet-of-things devices using pareto-optimal artificial neural network and multiple linear regression. IEEE Access 8:39525–39535
4. Mansour M, Sultan K, Kanaya H (2020) Compact dual-band tapered open-ended slot-loop antenna for energy harvesting systems. Electron 9(9):1394
5. Sultan KS, Abdullah HH, Abdallah EA, El-Hennawy HS (2020) Metasurface-based dual polarized MIMO antenna for 5G smartphones using CMA. IEEE Access 8:37250–37264
6. El Atrash M, Abdalla MA, Elhennawy HM (2019) A wearable dual-band low profile high gain low SAR antenna AMC-backed for WBAN applications. IEEE Trans Antennas Propag 67(10):6378–6388
7. Adhiyoga YG, Apriono C, Rahardjo ET (2020) Radiation characteristics of artificial magneto-dielectric antenna loaded with split-ring resonator array. Adv Electromagnet 9(3):32–38
8. Shahgholi A, Moradi G, Abdipour A (2020) Low-profile frequency-reconfigurable LTE-CRLH antenna for smartphones. IEEE Access 8:26487–26494
9. Van Trinh T, Jung CW (2019) Compact broadband internal monopole antenna with parasitic strips and sleeve feed for UHD-TV applications. IET Microwaves Antennas Propag 13(12):2096–2101
10. Navarro A, Mostardinha P, Varum T, Matos J, Maslovski S (2020) Double-dielectric microstrip ultrahigh-frequency antenna for digital terrestrial television. Appl Sci 10(23):8640
11. Ali SM, Sovuthy C, Imran MA, Socheatra S, Abbasi QH, Abidin ZZ (2020) Recent advances of wearable antennas in materials, fabrication methods, designs, and their applications: state-of-the-art. Micromachines 11(10):888
12. Kim JP, Kim GH, Seong WM, Choi JH (2009) A tunable internal antenna with an epsilon negative zeroth order resonator for DVB-H service. IEEE Trans Antennas Propag 57(12):4014–4017
13. Liu L, Rigelsford J, Langley R (2013) Tunable multiband handset antenna operating at VHF and UHF bands. IEEE Trans Antenna Propag 61(7):3790–3796
14. Zuo SL, Zhang ZY, Yang JW (2013) Planar meander monopole antenna with parasitic strips and sleeve feed for DVB-H/LTE/GSM850/900 operation in the mobile phone. IEEE Antenna Wireless Propag Lett 12:27–30
15. Anguera J, Puente C, Andújar A, Mateos RM, Vye D, Lien M (2021) Antenna library for IoT devices with antenna boosters. In: 2020 50th European microwave conference (EuMC), (pp 1147–1150), IEEE
16. Yeo J, Lee JI (2021) Miniaturized wideband loop antenna using a multiple half-circular-ring-based loop structure and horizontal slits for terrestrial DTV and UHD TV applications. Sens 21(9):2916
17. IEEE C95.1–2005 (2005) IEEE standards for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 kHz to 300GHz. Institute of electrical and electronics engineers, New York, NY

18. El Atrash M, Abdalgalil OF, Mahmoud IS, Abdalla MA, Zahran SR (2020) Wearable high gain low SAR antenna loaded with backed all-textile EBG for WBAN applications. IET Microwaves Antennas Propag 14(8):791–799
19. CST Microwave Studio Suite (2011) User's manual, www.cst.com
20. Fornes-Leal A, Cardona N, Frasson M, Castelló-Palacios S, Nevárez A, Beltrán VP, Garcia-Pardo C (2019) Dielectric characterization of in vivo abdominal and thoracic tissues in the 0.5–26.5 GHz frequency band for wireless body area networks. IEEE Access 7:31854–31864
21. Raveendran A, Sebastian MT, Raman S (2019) Applications of microwave materials: a review. J Electron Mater 48(5):2601–2634
22. Akila D et al (2021) In: Journal of physics: conference series 1963 012144
23. Regin R, Obaid AJ, Alenezi A, Arslan F, Gupta AK, Kadhim KH (2021) Node replacement based energy optimization using enhanced salp swarm algorithm (Es2a) in wireless sensor networks. J Eng Sci Technol 16(3):2487–2501

Comparative Analysis of KNN Classifier with K-Fold Cross-Validation in Acoustic-Based Gender Recognition

Disha Handa and Kajal Rai

Abstract Gender recognition based on acoustic attributes plays an important role in various audio forensic level tasks. When we talk about forensic level issues, accuracy is the most prominent attribute that needs to be taken care of. This article shows our attempts to observe the impact of multiple folds applied to the popular KNN classifier on the accuracy of results while recognizing the gender of the speaker. We demonstrate our experiments by using python programming language on the dataset available on Kaggle. The results show that 20-folds KNN can provide maximum accuracy (95.77%) and saturate afterward with the size of the dataset up to 3168. Results also show that changing the number of nearest neighbors in this algorithm will not put any impact on the accuracy.

Keywords Machine learning · Gender recognition · K-fold cross-validation · KNN classifier · Comparative analysis · Audio forensic

1 Introduction

Several potential applications can be derived from the automatic recognition of gender. Since global security is currently a concern, gender classification has attracted a great deal of attention, especially among speech researchers. There are several acoustic characteristics of a speaker that can help classify its gender. Features such as Short-Term Energy, Sound Intensity, Mel-Frequency Cepstral Coefficients, Pitch, and Frequency can differentiate male and female speakers as well as can identify gender-specific distress or joyful non-speech patterns [1, 2]. Many researchers have also done remarkable work in the area of gender recognition by using numerous AI, ML, and hybrid classification techniques [3]. Besides all this significant work, a need for understanding the accuracy level that the classifier can provide is required. Among the three datasets, few claim they have achieved 96.8% accurateness with KNN for the TIMIT dataset [4].

D. Handa (✉) · K. Rai
University Institute of Computing, Chandigarh University, Mohali, India
e-mail: dishah@gmail.com

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_34

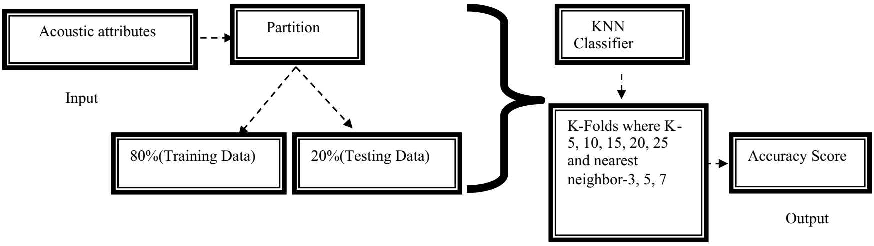


Fig. 1 Block diagram for gender identification

Similarly, every researcher has claimed the effectiveness of the algorithm or classifier applied. In this study, we compare the accuracy level of a single classifier, KNN, using k-folds on the same dataset. As a result, we learn how many folds one can apply to achieve maximum accuracy. Figure 1 depicts the model we have used for the study. Python is well known for its flexibility and its vigorous community. Therefore, working with machine learning algorithms or classifiers is not quite as complex as developing the entire algorithm from scratch. This study relies on Matplotlib, Pandas, NumPy libraries, Scikit, sklearn, kNeighborsRegressor, and neighbors Classifier.

1.1 Literature Review

In both classification and regression, KNN algorithm is used. The weighted averages are calculated by assigning neighbors' contributions relative to their distance from each other. Across all parameters of consideration, the KNN algorithm performs well. The reason it is used most often is because it is easy to interpret and requires less computation. Several researchers have employed k-fold cross-validation in a wide range of experiments [5–8]. In [9], authors proposes a computer-aided diagnosis approach based on KNN with k-fold cross-validation and neural network technique. With machine learning (ML) methods, k-fold cross-validation, and performance evaluation metrics, the researchers were able to achieve an average accuracy of 96.78% with the initial training data. In [10], authors introduce six types of machine learning algorithms based on similar characteristics, like KNN, decision trees, Naïve Bayes, Linear Regression, and SVM. This study evaluates machine learning algorithms, which are tested for accuracy. The study uses the k-fold cross-validation algorithm as a model for evaluating the effectiveness of machine learning algorithms in data analytics. There is a 95% accuracy rate for fold-3, which is the best test for accuracy.

2 Method and Experiment

We have used K-Nearest Neighbor (KNN) Classifier for classification of acoustic data to distinguish gender of the speaker. This classifier is popularly used to differentiate the data into different classes based on various attributes. It comes under the category of supervised machine learning algorithms and one of the widely used machine learning algorithms [12]. The entire process is divided into the following tasks:

A. *Importing and Reading Data*
To implement a KNN classifier using Python, we have used a dataset available on Kaggle. The file has 3168 rows and 21 columns. Data is shown in the first 20 columns, and a label appears in the final column that represents gender [13]. There are worth twenty acoustic attributes such as mean free, sd, and median. In the first step, we have to import and read the dataset.

B. *Preprocessing*
In general, AI and machine learning algorithms perform better when given numerical input. So, it is considered a challenge to change untrained data into mathematical form but at the same time maintain its efficiency. There are several methods to change data into numeric format. Trade-offs and effects of various techniques are different on different feature sets [14, 15]. In our experiment, we use the Label-Encoder method to convert categorical data into numerical form by using the SciKit-learn library. The next step is to perform normalization to ensure that all the data lies in a similar range.

C. *Constructing Model using KNN Classifier*
It is necessary to divide the data into two categories: training and testing. We classified 80% of the data as training data and 20% as testing data in our study. Afterward, we have constructed a regression model of KNN with 3, 5, and 7 nearest neighbors.

D. *Evaluation*
With certain types of data, cross-validation works very well for evaluating machine learning models. The method uses a particular parameter, k, which represents the number of groups that a specific data set is to be broken into. Hence, it is known as k-fold cross-validation where k can be any positive integer number greater than zero. This technique is mainly used in applied ML to estimate the ability of any model on anonymous data. The dataset is shuffled randomly in this approach and then split into k groups. The model is fitted on a training set, evaluated on a test set, and evaluated on a test set using one group as the test data set and the rest as training data sets. Each of the k subsamples is used as the validation data exactly once during the cross-validation process. An average can then be calculated based on the results of k measurements of accuracy. The examinations in the data set are assigned to groups and remain in those groups during the entire procedure. Further, each sample is allowed to appear in the hold-out set only once and to be used to train the model k times [16, 17]. Figure 2 shows how to do cross-validation using 10-folds in Python; the accuracy score is also displayed in percentage that is 95%.

```
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import cross_val_score
knn_scores = []
for k in range(1,21):
    knn_classifier = KNeighborsClassifier(n_neighbors = k)
    score = cross_val_score(knn_classifier,X_train,y_train,cv=10)
    knn_scores.append(score.mean())

s52 =score.mean()
print(s52)
```

```
0.955440233606297
```

Fig. 2 K-fold cross-validation using KNN

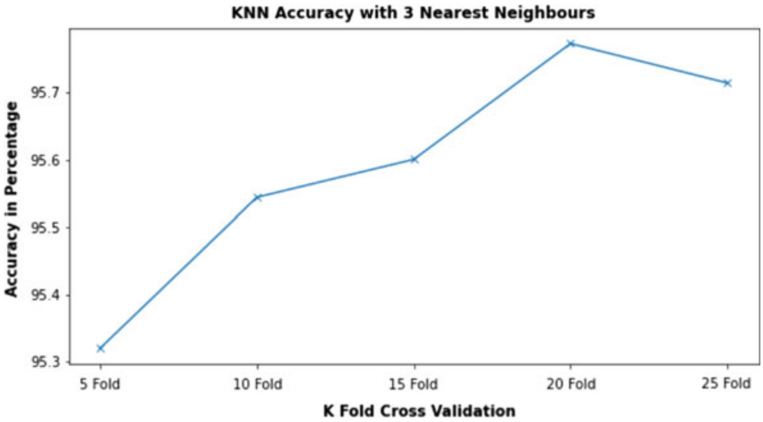


Fig. 3 K-fold cross-validation with 3 nearest neighbors

We have built a KNN classifier using 5-, 10-, 15-, 20-, and 25-folds of cross-validation to make a comparison based on the accuracy provided by the specific fold. Figure 3 shows the accuracy scores of KNN using 5–25 k-folds with 3 nearest neighbors. It is clear from that with the increase of the folds the accuracy rate increases and after 20-folds it starts degrading for this particular dataset. It further states that up to 3000 set of records, we can apply 20-fold to achieve maximum accuracy. The similar type of experiment is being done with 5 nearest neighbors to examine its impact. Figure 4 depicts the KNN accuracy rate with 5 nearest neighbors. Table 1 reflects that the increase of number of neighbors is not putting any impact on the result set.

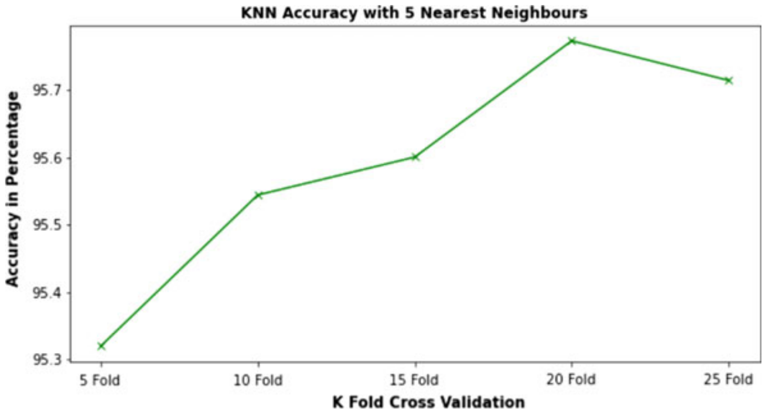


Fig. 4 K-fold cross-validation with 5 nearest neighbors

Table 1 KNN model accuracy with k-folds

No. of nearest neighbor	Accuracy in percentage (%)				
	5-Folds	10-Folds	15-Folds	20-Folds	25-Folds
3	95.32	95.54	95.6	95.77	95.71
5	95.32	95.54	95.6	95.77	95.71
7	95.32	95.54	95.6	95.77	95.71

3 Conclusions and Future Work

With the help of a widely used KNN classifier, the study examined acoustic attributes and classified them into male and female categories. The KNN classifier has been used for gender detection by many researchers. This study, however, is primarily concerned with the accuracy that the KNN algorithm can provide with multiple folds and for *n* neighboring points where *n* can be any positive number. The agenda is to find the threshold value of accuracy with multiple folds and *n* neighbors'. The aim is to find the threshold value of accuracy with multiple folds and *n* neighbors'. It has been found that the KNN with 20-folds provides maximum accuracy and starts to degrade afterward. On this dataset, however, the number of neighbors does not affect the accuracy. In future studies, we will examine more attributes in a given dataset that can impact accuracy.

References

1. Pahwa A, Aggarwal G (2016) Speech feature extraction for gender recognition. Int J Image Graphics Sign Process 8(9):17
2. Handa D, Vig R (2020) Distress screaming vs joyful screaming: an experimental analysis on both the high pitch acoustic signals to trace differences and similarities. In: 2020 Indo–Taiwan 2nd international conference on computing, analytics and networks (Indo-Taiwan ICAN), (pp 190–193), IEEE
3. Ertam F (2019) An effective gender recognition approach using voice data via deeper LSTM networks. Appl Acoust 156:351–358
4. Uddin MA, Hossain MS, Pathan RK, Biswas M (2020) Gender recognition from human voice using multi-layer architecture. In: 2020 international conference on innovations in intelligent systems and applications (INISTA), (pp 1–7), IEEE
5. Gadekallu TR, Rajput DS, Reddy M, Lakshmanna K, Bhattacharya S, Singh S, Jolfaei A, Alazab MA (2021) A novel PCA–whale optimization-based deep neural network model for classification of tomato plant diseases using GPU. J Real-Time Image Process 18(4):1383–1396
6. Basha SM, Rajput DS, Iyengar NCSN, Caytiles DR (2018) A novel approach to perform analysis and prediction on breast cancer dataset using R. Int J Grid Distrib Comput 11(2):41–54
7. Kolisetty VV, Rajput DS (2020) A review on the significance of machine learning for data analysis in big data. Jordanian J Comput Inf Technol (JJCIT), 6(01)
8. Basha SM, Rajput DS (2018) Parsing based sarcasm detection from literal language in tweets. Recent Pat Comput Sci 11(1):62–69
9. Cherradi B, Terrada O, Ouhmida A, Hamida S, Raihani A, Bouattane O (2021) Computer-aided diagnosis system for early prediction of atherosclerosis using machine learning and K-fold cross-validation. In: 2021 international congress of advanced technology and engineering (ICOTEN), (pp 1–9), IEEE
10. Tamilarasi P, Rani RU (2020) Diagnosis of crime rate against women using k-fold cross validation through machine learning. In: 2020 fourth international conference on computing methodologies and communication (ICCMC), (pp 1034–1038), IEEE
11. Tembusai ZR, Mawengkang H, Zarlis M (2021) K-nearest neighbor with k-fold cross validation and analytic hierarchy process on data classification. Int J Adv Data Inf Syst 2(1):1–8
12. Fix E, Hodges JLFix E, Hodges JL (1989) Discriminatory analysis. Nonparametric discrimination: Consistency properties. Int Stat Rev 57(3):238–247
13. https://www.kaggle.com/primaryobjects/voicegender
14. Potdar K, Pardawala TS, Pai CD (2017) A comparative study of categorical variable encoding techniques for neural network classifiers. Int J Comput Appl 175(4):7–9
15. Hancock JT, Khoshgoftaar TM (2020) Survey on categorical data for neural networks. J Big Data 7(1):1–41
16. Stone M (1974) Cross-validatory choice and assessment of statistical predictions. J Roy Stat Soc: Ser B (Methodol) 36(2):111–133
17. Stone M (1977) An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. J Roy Stat Soc: Ser B (Methodol) 39(1):44–47

Modified ElGamal Algorithm Using Three Paring Functions

Eman Hatem Omran and Rana Jumaa Sarih Al-Janabi

Abstract Cryptography defines different methods and technologies used in ensuring that communication between two parties over any communication medium is secure, especially in presence of a *third* part. This is achieved through the use of several methods, such as encryption, decryption, signing, generating of pseudo-random numbers, among many others. Cryptography uses a key or some sort of a password to either encrypt or decrypt a message that needs to be kept secret. This is made possible using two classes of key-based encryption and decryption algorithms, namely symmetric and asymmetric algorithms. The best known and the most widely used public key system is ElGamal. This algorithm comprises of three phases, which are the key generation phase, encryption phase, and the decryption phase. Owing to the advancement in computing technology, ElGamal is prone to some security risks, which makes it less secure. The following paper previews combination of three paring function used to enhance the ElGamal algorithm and increase its security. The results showed that the modified algorithm gives 93% accuracy.

Keywords Cryptography · ElGamal algorithm · Encryption · Decryption · Cryptosystem · Security · Public key · Private key · Paring functions

1 Introduction

Cryptography is the study of algorithms that provides a security service and protects the integrity of data, algorithms that guarantee the authenticity of the source of data, and algorithms that provide confidentiality for data (encryption algorithms). Privacy is at the heart of cryptography. Coding is a functional means to achieve data privacy [1]. Cryptography played a vital role in many aspects of our world today, such as online banking and E-commerce operations and E-mail. Understanding the principles

E. H. Omran (✉) · R. J. S. Al-Janabi
Computer Science Department, University of Alqadisiyah, qadisiyah, Iraq
e-mail: com.post15@qu.edu.iq

R. J. S. Al-Janabi
e-mail: rana.aljanaby@qu.edu.iq; rana.aljanaby@gmail.com

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_35

of encryption depends on the knowledge of many topics such as complexity and pure mathematics [2]. Encryption is separated into two basic forms (symmetric encryption also named private key encryption uses the similar key to encode and decode like stream cipher and block cipher), and asymmetric encryption also named public key coding uses two couples of key, one to code the message and the other to the encode, the first is identified as public key because it is identified to operator in the chosen situation, it is utilized to code messages, while the next is known as private key so termed because he is recognized to only one operator which is the owner and is utilized to decode encrypted messages public key. The goals of cryptography can be descripted as the following:

(a) **Confidentiality**: It confirms that only official personnel can access data. Hiding data by encrypting information is one method to provide secrecy.
(b) **Data integrity**: It confirms that it is possible to identify unauthorized adjustments to the information. It also saves against an attacker changes documents through transportation, like interrupting an E-mail message and changing the message before sending it to the receiver.
(c) **Authentication**: It confirms that data is created from the operator or computer that claims to have sent the data. It also saves against imposters and man-in-the-central attacks.
(d) **Non-repudiation**: It certifies that an operator cannot deny acting a job or transfer data. For sample, non-repudiation confirms that a party to an agreement cannot refute having signed the agreement.
(e) **Anti-replay protection**: Avoids an attacker from interrupting a message and transfers it later. For example, an attacker can detention a logon sequence and then replay the system packs to logon at a later time. Anti-replay protections, like addition encrypted time brands to information, stop such attack.

2 Related Works

Galindo and Großschädl [3], they said that leakage-flexible cryptography purposes to expand the rigorous promises succeeded over the verifiable safety example to physical applications. The structures planned on based of this different method necessarily hurt from an Achilles foot: A restricted outflow supposition is required. At present, a massive gap occurs amid the concept of many projects and their application to approve the outflow flexibility in training. The current job efforts to thin this hole for the outflow-tough interconnected ElGamal key encapsulation device (BEG-KEM) offered in 2010. Their main influence in the different of a restricted seep and first-calculation-seepages ideal that is nearer to training. We deteriorate the limit on the document extent of the leak jobs in all types.

Siahaan and Elviwani [4], an asymmetric process is a coding method that utilizes alternative solutions on the procedure of coding and decoding. This process utilizes two answers, open key and special key. The open key is openly divided, whereas the user keeps the special key privately, and answer needs the period for decoding procedure. RSA and ElGamal are dual processes that appliance an open key cryptosystem. The power of the process dishonesties in the while distance utilized. The gradation of struggle in RSA is located in the separating of big peaks, whereas ElGamal is located in a computation of separate algorithms. In later experimentation, it is confirmed that RSA completes a quicker coding procedure than ElGamal, while ElGamal decoding procedure is quicker than RSA. All of these processes are cryptographic open key processes, then got purposes in various techniques. RSA is an imperative process, whereas ElGamal is a potential process.

Magsino and Arboreta [5], they supposed that the credit card amount could be protected by hiding the main figures to a ciphertext. Altered techniques of encryption could be utilized, but several of individuals are disposed to each physical strength offense particularly and have been utilized through various. This suggests into mixture of the ElGamal coding system and RSA and chaos process. The originality of scheme and haste has been proved to show the competence of the fresh scheme. The tests have proved that the algorithm of the new cryptosystem is more secure than its parent cryptosystems (RSA and ElGamal). But, the speed of the new system is slightly slower than its parents. The RSA algorithm can be slower depending on the chosen encryption key. Overall, the new cryptosystem is found efficient to use in credit card number encryption.

Jia1 et al. [6] ElGamal cryptography is unique of the greatest significant public key cryptography (PKC), meanwhile it was suggested while these PKCs which are grounded on solid issue that separate process issue and numeral factorization issue are weak with improvements in volume CPUs. So, selected replacements must be suggested two ElGamal-like open key coding systems grounded on big abelian subcategory of official linear collection over a remainder ring, but the two structures were not extended; earlier, it was verified risky through us. Then, in (2016), they projected a better-quality cryptosystem, which contains resistance of my occurrence on ‘NEURAL COMPUTING & APPLICATIONS.’ Through examining the safety of the open key cryptography, we suggest an enhanced technique of arithmetical answer-resuscitation offense in the mathematical calculation difficulty despising the inventors’ right that the cryptosystem is best safety. In addition, they deliver consistent applied attack instance to show the offense process in our cryptanalysis, whereas disruptions examples are demanding 10 bytes of safety fewer than 60 s on a solo PC workstation.

Mani and Begam [7], in their paper, they supposed that the possible weakness of ElGamal cryptosystem is the ciphertext shaped which is continuously doubled as extended to the normal text to the communication development via a feature of dual earnings location through coding. After the letter is very lengthy, the ciphertext formed by the ElGamal cryptosystem is also extended for example; once the ciphertexts are conveyed over the message station, which goes to deliver fewer safeties cause, if the opponent interrupts anybody of the ciphertext after dual ciphertexts for both charm of the Normal text, the additional might be saved simply because there is a connection in the middle of the dual ciphertexts. Doubt dual ciphertexts are descending by one; the opponent might not be talented to expect the dual ciphertexts from single. To improve the safety of ElGamal algorithm, the dual Cantor purpose, Rosenberg pairing purpose, and Elegant pairing purposes are utilized in this research. Once the supposed meanings are utilized, the dual ciphertexts shaped through each normal text charm are decline by one, so that the opponent will not simply be improved by the normal text. New outcomes obviously exposed improving the safety of ElGamal algorithm afterward joining the combination jobs in it.

3 Asymmetric Cryptography

Once it arrives to the term 'Encryption,' we celebrate it as a method that guards documents employing a cryptographic answer, and there is nothing incorrect with this. Nevertheless, what greatest people do not understand is that there are a lot of types of coding systems. Asymmetric coding, also named as public key cryptography, is an illustration of one kind. Different 'standard' (symmetric) coding, asymmetric encryption code and decode the information by dual isolated exactly linked cryptographic answers. These answers are called as a 'Public Key' and a 'Private Key.' Both are named as 'Public and Private Key pair. Let us show how these dual keys action with each other to make the difficult power that is asymmetric coding [8].

Asymmetric coding usages dual different, yet connected keys. First key, the open key, is applied for coding, and the second, the special Key, is for decoding. As indirect in the term, the private answer is proposed to be secret so that just the certified receiver could decode the letter.

Let us know this with a asymmetric encryption sample. Imagine you are a snooping action and you want to plan an apparatus for your managers to transfer it safely. You do not want double-method statement, they got their instructions, and you only want normal itemized reports upcoming in from them. Asymmetric coding would let you to make public answer for the operator to code their data and a private answer back at control center that is the just method to decode it all. This offers a solid system of first-technique connection (Fig. 1).

At the main of asymmetric coding drops a cryptographic process. This procedure uses a main group procedure (a type of scientific purpose) to produce a key pair. All these keys are arithmetically related with one another. This connection in these keys is different from one system to other. The process is a mixture of dual jobs—coding purpose and decoding purpose. To state the clear, the coding mean codes the information and decodes meaning decode it [9].

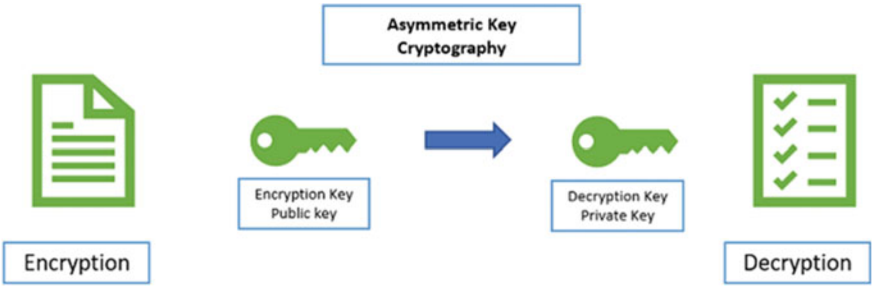


Fig. 1 Asymmetric key cryptography

4 Proposed System

The proposed system in this paper is to encrypt and decrypt the texts by adding three paring functions to ElGamal algorithm. Their paring functions are Cantor pairing function, Rosenberg-Strong pairing algorithm, and Elegant pairing algorithm, the goal of this algorithm is to check if their functions are getting better encryption and decryption results than the original ElGamal algorithm, and the steps of the proposed algorithm are showed in Algorithm 1 and in Fig. 2.

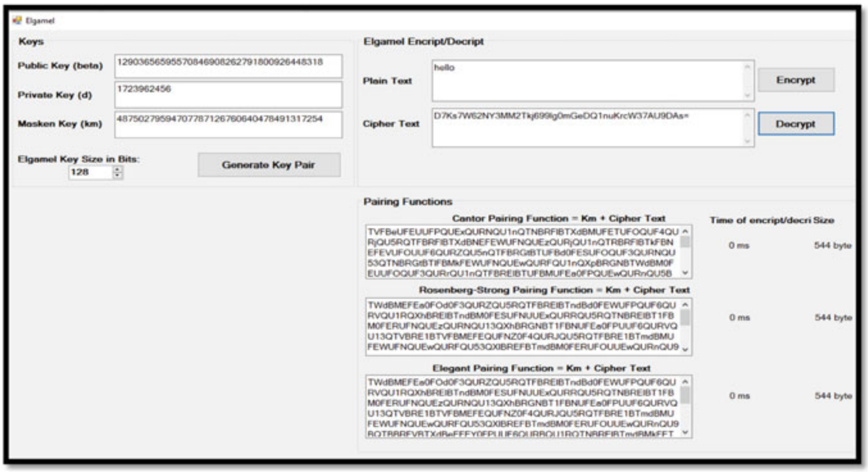


Fig. 2 Execution of the combination algorithms

Algorithm 1: the combination of ElGamal with 3 paring functions
Input: plaintext
Output: ciphertext (encryption and decryption)
Begin **Step1**: generate random keys for cipher in ElGamal and for that we use this method to generate large primary random number for the key (p) **Step2**: generate other random number by bit number **Step3**: start cipher using ElGamal algorithm **Step4**: start cipher using main method named Cantor Pairing Function **Step5**: now cipher using next method named Rosenberg-Strong Pairing Function **Step6**: Start cipher using third method named Elegant Pairing Function **Step7**: display the result of ciphering on monitor **Step8:** finally decrypt the ciphertext to get the plaintext once again **End**

(a) Cantor Pairing Function
Assume some group B, a pairing purpose for B is a 1–1 message of the group of orderly pairs B^2 to the group B. The usual B is supposed to remain limited with pairing purposes; it consumes less from dual features. A pairing purpose aimed at B essentially occurs, if B is unlimited. The Cantor's pairing purpose [10, 11] to the numbers is of the method

$$c(x, y) = z = 1/2(x^2 + 2xy + y^2 - x - 3y + 2) \tag{1}$$

(b) Rosenberg-Strong Pairing Function
The Rosenberg-Strong pairing algorithm [12] to the undesirable numbers is clear via the formulation

$$r(x, y) = (\max(x, y))^2 + \max(x + y) + x - y \tag{2}$$

In the setting of the Rosenberg-Strong pairing function, the amount max(x, y) is supposed to be the seashells integer of the point (x, y). The opposite of the Rosenberg-Strong combination purpose r(x, y) is specified by the formulation

$$r^{-1} = \begin{cases} (z - m^2, m) & , if\ z - m^2 < m \\ (m, m^2 + 2m - z) , & \text{otherwise} \end{cases}$$

(c) Elegant Pairing Function
If x and y are undesirable numbers of Elegant pairing algorithm. Formerly, E(x, y) outcome is only undesirable number that is exclusively related with that pair [13].

$$E(x, y) = z = \begin{cases} y^2 + x & x \neq \max(x, y) \\ x^2 + x + y & x = \max(x, y) \end{cases}$$

Table 1 ElGamal algorithm enhanced with pairing functions experimental results of ten times the sentence for 128 key sizes

Algorithm name	Encryption time (MS)	Decryption time (MS)	Total time (MS)	Encryption speed (kb/s)	Decryption speed (kb/s)	Total speed (kb/s)	Block size (bits)
ElGamal (default)	0.001	0.01	0.011	1648.43	164.84	1813.27	128
Cantor	0.005	0.031	0.036	4227.343	681.82	4909.163	128
Rosenberg-Strong	0.003	0.025	0.028	7045.57	845.46	7891.03	128
Elegant	0.009	0.074	0.083	2348.52	285.63	2634.15	128

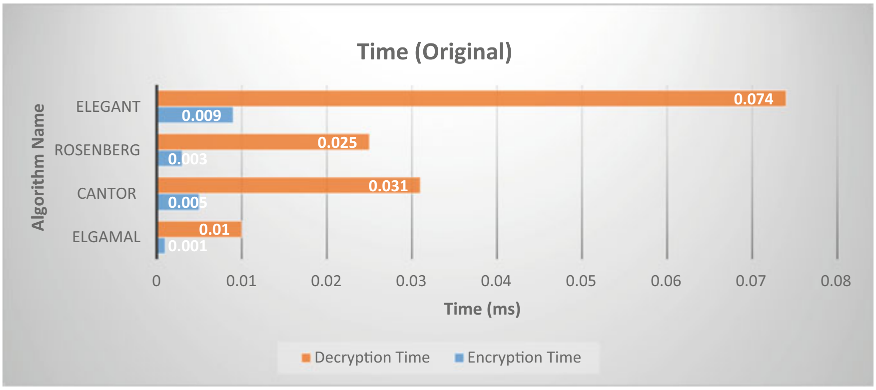


Fig. 3 Encryption and decryption time across the proposed algorithms for 128-bit key size

5 Experimental Results

This is the enhanced version of our proposed system where a plaintext goes through two ciphering operations; the first one is the default ElGamal algorithm, while the second encryption is through one of the pairing functions (Cantor, Rosenberg-Strong, and Elegant). Hence, there are three ciphertexts. The experimental results of ciphering ten times the sentence '**the quick brown fox jumps over the lazy dog**' are shown in Table 1 and are for 128 key sizes in bits, respectively, (Figs. 3 and 4).

6 Conclusion

An improved form of ElGamal coding with triple alternative pairing algorithm is supposed of and applied them effectively. The trial outcomes obviously designated to rise in the safety rank of ElGamal coding when pairing algorithm are applied into it. The purposes of getting the right evaluation results of applying the plaintext on different algorithms giving ElGamal key size a fixed value each time makes it easier

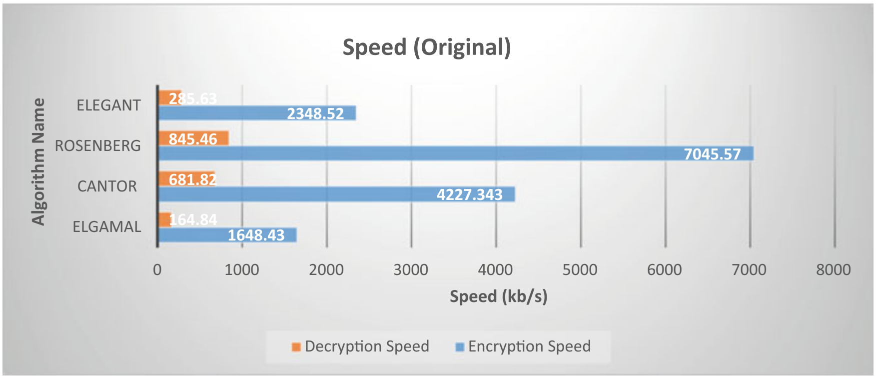


Fig. 4 Encryption and decryption speed across the proposed algorithms for 128-bit key size

to monitor the change in encryption measurement parameters as shown in Table 1 above by setting the key size to 128. The results show that with a fixed block size and key size, the encryption and decryption speed of Cantor pairing function is faster than ElGamal default algorithm and both Rosenberg-Strong and Elegant pairing functions. With that is said, the encryption and decryption time of both Rosenberg-Strong and Elegant pairing functions is less than others. On the other hand, the ciphertext size of all pairing functions is greater that ElGamal default algorithm in multiple times. Of course, all of that has to do with how each algorithm works especially that each algorithm has different number of mathematical operations.

References

1. Dent AW, Mitchell CJ (2005) User's guide to cryptography and standards, Artech House. INC, USA
2. Talbot J, Welsh D (2006) Complexity and cryptography an introduction. Cambridge University Press, UK
3. Galindo D et al (2014) Implementation of a leakage-resilient ElGamal key encapsulation mechanism Springer special section on proofs
4. Siahaan APU, Elviwani, Octavian B (2018) Comparative analysis of RSA and elgamal cryptographic public-key algorithms EAI
5. Magsino JP, Arboleda ER, Corpuz RR (2019) Enhancing security of El Gamal encryption scheme using RSA and chaos algorithm for E-commerce application. Int J Sci Technol Res
6. Jia J, Wang H, Zhang H (2019) Cryptanalysis of an ElGamal-like cryptosystem based on matrices over group rings, Springer Nature Singapore Pte Ltd
7. Mani K, Begam BA (2020) Enhancing the security in ElGamal cryptosystem using paring functions. Int J Innov Technol Exploring Eng (IJITEE), 9(4)
8. Tarnish AH (2000) Designing and implementing a stream cipher image cryptography system. (M.Sc. Thesis, University of Technology at Computer Science)
9. Kiraz MS, Uzunkol O (2018) Still wrong use of pairings in cryptography. Appl Math Comput, Elsevier 333:467–479
10. Mrabet MJ (2017) Guide to pairing-based cryptography, 1st edn. Chapman and Hall/CRC, New York

11. Cegielski P, Richard D (2001) Decidability of the theory of the natural integers with the cantor pairing function and the successor. Theor Comput Sci, Elsevier 57:51–77
12. Szudzik MP (2019) The rosenberg-strong pairing function. Discrete mathematics
13. Matthew S (2006) An elegant pairing function. Wolfram Research, Inc.

Mental Health Analysis and Classification During Covid-19 Using Big Data Approach

Bhanvi Badyal, Hrishabh Digaari, and Tarun Jain

Abstract In December 2019, a deadly virus named SARS-CoV-2 started spreading in the regions of Wuhan, Hubei, China. The number of coronavirus patients gradually increased in Wuhan, and by 20 December, it reached 60 and 266 by 31 December. Till now, there have been more than 40 Lakhs deaths due to Covid-19. This deadly pandemic gave a setback to most people all over the world in terms of losing their loved ones. Apart from that, this pandemic mentally affected a lot of minds. Social illness and loneliness have been linked to poor mental health by a broad body of research, and data from late March suggests a negative increase in mental health. There had been news of people committing suicides or some going under depression all because their social life was cut down and all they did was question their life choices, their existence, their personality, and their achievements which ultimately trapped them in those intrusive thoughts that kept popping up again and again—which made them disturbed or even distressed. The objective of this paper is to analyze and categorize the mental states of people from all over the world in order to raise mental health awareness, particularly during COVID-19. We used the big data approach to display the surge in sadness and suicidal ideation in terms of the increase in the frequency of certain words. To continue with this problem statement, we will examine text data and learn what words are utilized in virtual suicide/depression notes utilizing two subreddits and NLP tools.

Keywords Covid-19 · Mental health analysis · Multinomial Naïve Bayes · NLP · Big data analysis

1 Introduction

Depression (major depressive disorder) is a widespread and significant medical condition that has a negative impact on how you feel, think, and behave [1]. It can be described as a feeling of sadness, loss, or rage that interferes with one's daily life.

B. Badyal · H. Digaari · T. Jain (✉)
SCIT, Manipal University Jaipur, Jaipur, India
e-mail: tarunjainjain02@gmail.com

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_36

In such circumstances, we wished we were there for the individual to listen to their grief. With the advent of online support forums like Reddit, it is easy for people to express themselves to receive help. At this particular time, they need only one ear to hear them.

A recent study of 1441 participants from the period of COVID pandemic and 5065 participants prior to [2] the pandemic showed that the prevalence of depression during the COVID-19 pandemic was more than three times higher than before. The WHO recommends that countries develop a Suicide Prevention Strategy that includes programs aimed at "early detection of groups at risk of suicide."

In today's world, social media provides individuals with a platform to express themselves and receive the help they require from their peers [3]. Individuals can communicate their ideas on their vulnerabilities and mental states in a safe and anonymous atmosphere on such social networks.

We took advantage of this chance to collect data from social media sites as a source to assess depressive thoughts and detect suicidal propensity in individuals. Such analysis also helps to reform our understanding of different adjustments in self-awareness and introspection.

1.1 Characteristics of Big Data: The 5Vs

Big data can be described by the following characteristic:

Volume. Displays a great quantity of data gathered from many sources

- The volume of data for this project scales up to 60,000 + posts from the active community of Redditors. Each post has between 30 and 35 features.
- The most suited form of storage is cloud.
- Raw data can be discarded once we have processed data with selected features.

Variety. Incoming data might have different types broadly categorized into semi-structured, structured, unstructured, and quasi-structured

- Only two subreddits with over one million active users were chosen for the scope of this study.
- Scraped quasi-structured data is present in the format of JSON.

Value. Data that is helpful from a large amount volume of data.

- There were originally 30–35 characteristics, but only 10 made it into the final classification model.

Velocity. Measures how quickly data can be acquired, analyzed, and used.

Veracity. The data's correctness, legitimacy, and accuracy.

- We scraped text post of real time made by supporters of particular social communities; therefore, the data sources are reliable [4].

2 Methodology

The big data life cycle has the following stages (Fig. 1).

2.1 *Discovery*

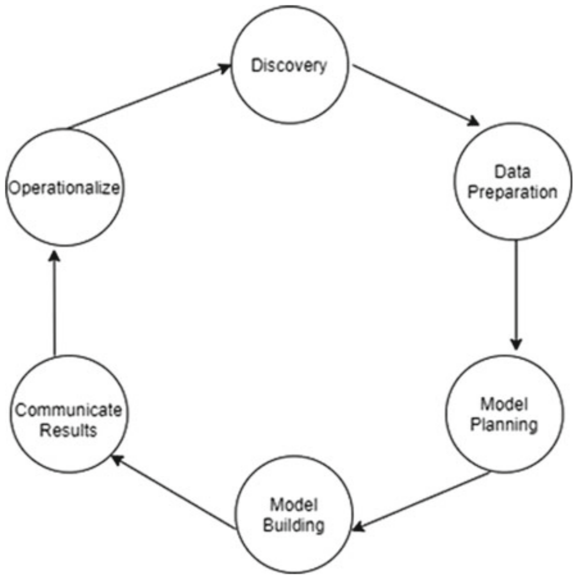
Once we decided to proceed with this topic, we began our groundwork by reading a lot of research papers and articles on the Internet. We observed that there were quite a few papers on mental health analysis but none of the papers studied the consequences of Covid-19 on the mental stress of the people [5–8]. So, using a similar approach but in a different scenario, we decided to analyze the social media posts and also build a classifier to distinguish between the mental states, to increase mental health awareness, via completing the other steps of the big data life cycle.

2.2 *Data Preparation*

We identified two subreddit support communities that provided us with enough information to identify those who are at risk of depression or suicidal ideation.

Below were the subreddits used:

Fig. 1 The big data life cycle



- r/depression—their idealism is that no one should be alone in the dark phase of their life (693 k members)
- r/SuicideWatch—they offer peer help to individuals who are struggling with suicidal thoughts (234 k members) [9].

We scraped data from many subreddits using the Reddit API. We scraped semi-structured material. It was in text format, but it was difficult to format. As a result, we created a function to transform the JSON data to a tabular representation. However, the analysis of data in its current form was not yet possible.

We procured around 60,000 entries which were then divided into 2 separate sets, each with around 100 columns. To distinguish between the two sources, we enhanced the dataset by an extra column labeled "is_suicide."

Another part of the data preparation included data cleaning [10]. Firstly, we had to select data that was useful for the analysis. Then, we had to clean and filter out all of the superfluous and inconsistent data.

Hence, we carried out the following activities:

Data Selection. Selection of pertinent data that was useful for analysis.

- Cutting down the dataset—Because both sets included 100 columns, we chose a handful that would be beneficial for our research, which includes "post," "author," "title," and "number of comments," as well as "URL."
- Concatenation—As we had already created an "is_suicide" column indicating which subreddit the posts are from, we concatenated both datasets together [11].

Data Pre-processing. Cleaning and filtering superfluous and inconsistent data.

- Built processing functions—This allowed us to lowercase the text, use lemmatization to downgrade comparable terms to the base word, remove punctuation, and separate the clean data into separate rate columns [12].

3 Model Selection

We then used a grid search with cross-validation to rank various classifier models such as k-nearest neighbors [13] and Multinomial Naive Bayes [14] before deciding on the best model for our data.

- Establishing a baseline score: Calculated the baseline score for the models to "outperform." In the context of the project, a baseline score is the probability of correctly predicting that all of our Reddit postings would come from the r/SuicideWatch subreddit.
- Choosing the most appropriate parameter: To find accuracy on different dataset, we created a multi-model function. On separate columns, we ran a combination of Count Vectorizer and Naive Bayes model and rated them. See Table 1 for more information. This aided us in selecting a column on which to base future models.

Table 1 The following table shows the score of various models on different columns

Column used (X)	Model	AUC score	Precision	Recall (sensitivity)	Train accuracy	Test accuracy	Baseline accuracy	Specificity	$F1$-score
Selftext	CountVec + MultinomialNB	0.66	0.63	0.63	0.9	0.63	0.52	0.61	0.63
Author	CountVec + MultinomialNB	0.62	0.75	0.58	1	0.58	0.52	0.99	0.5
Title	CountVec + MultinomialNB	0.66	0.61	0.61	0.83	0.61	0.52	0.59	0.61
Selftext_clean	CountVec + MultinomialNB	0.67	0.64	0.64	0.9	0.64	0.52	0.6	0.64
Author_clean	CountVec + MultinomialNB	0.58	0.57	0.56	0.96	0.56	0.52	0.71	0.56
Title_clean	CountVec + MultinomialNB	0.66	0.63	0.63	0.83	0.63	0.52	0.63	0.63
Megatext_clean	CountVec + MultinomialNB	0.71	0.68	0.68	0.95	0.68	0.52	0.66	0.68

We proceeded with the megatext_clean column based on a mix of scores from the modeling exercise.

Some of the reasons behind this decision were as follows:

- Good Score: The test set of the model utilizing megatext_clean scored a 68% on the metric of f1-score, while the training set scored a 95% on the same.
- High ROC: AUC score is considered as an appropriate metric to quantify the quality of model predictions as the classes are essentially balanced. The AUC score of megatext_clean is the highest.
- True Positive Rate: For the megatext_clean metric, the model fared best.

Now, inspired by our earlier function, we created a similar function that ran multiple combinations of models with Count Vectorizer (cvec), Hashing Vectorizer (hvec), and TF-IDF Vectorizer (tvec) on the megatext_clean column. See Table 2.

4 Experimental Results

Based on a combination of results from the preceding modeling exercise, we proceeded with the TF-IDF and Multinomial Naive Bayes model. Although, the Hashing model had a slightly better AUC score, we preferred the TF-IDF + Multinomial Naive Bayes model because of the higher Recall score. This model also fits pretty well with only a 0.02 difference between its train and test set scores. Hence, the final production model was a combination of two models: TF-IDF and Multinomial Naive Bayes [15]. TF-IDF was invented to make the keywords rank low, in case they appear many times in a document. Since there are several overlapping keywords between our 2 classes, TF-IDF is expected to perform good.

To find how the covid situation was impacting the people worldwide, we did an analysis between the number of posts and the time period (2018–2020) (see Fig. 2). Our analysis found a significant increase in the posts during the year 2020. There may be many factors for this increase but according to our hypothesis, covid could be one of them.

Justifying the above statement, we visualized the most used words by comparing the Reddit posts before and during covid in a bar plot (Figs. 3 and 4) and a word cloud (Fig. 5). Word cloud is a data visualization tool for visualizing text data in which magnitude of each word represents its frequency or relevance. For forming the word cloud, we used a certain set of posts to avoid the extra computation problem. Those posts were sampled out using random sampling. A simple random sample (srs) is a subset of individuals chosen randomly, all with the same probability, from a larger set. In srs, each subset of k individuals has the same chance as every other subset of k persons of being chosen for the sample.

Below are the results from our analysis which differentiates the words used before and after covid. We discovered that, pre-covid period (2019), majority of the keywords used were “feel,” “depression,” “want,” “life,” and “suicide” but during covid period (2020), we observed that new keywords related to Covid-19 such as

Table 2 The following table shows the performance metrics for all the candidate models on megatext_clean column

Model	AUC score	Precision	Recall (sensitivity)	Best score	Train accuracy	Test accuracy	Baseline accuracy	Specificity	$F1$-score
cvec + multi_nb	0.71	0.65	0.65	0.65	0.69	0.65	0.52	0.65	0.65
cvec + ss + knn	0.63	0.59	0.59	0.59	0.73	0.59	0.52	0.56	0.59
cvec + ss + logreg	0.72	0.66	0.66	0.64	0.68	0.66	0.52	0.7	0.66
tvec + multi_nb	0.74	0.67	0.66	0.65	0.69	0.66	0.52	0.73	0.66
tvec + ss + knn	0.63	0.59	0.59	0.6	0.74	0.59	0.52	0.64	0.59
tvec + ss + logreg	0.74	0.67	0.67	0.64	0.7	0.67	0.52	0.7	0.67
hvec + multi_nb	0.74	0.69	0.67	0.69	0.87	0.67	0.52	0.83	0.66
hvec + ss + knn	0.57	0.5	0.5	0.5	0.64	0.5	0.52	0.02	0.35
hvec + ss + logreg	0.68	0.64	0.64	0.65	1	0.64	0.52	0.55	0.63
hvec + multi_nb (tuning)	0.71	0.64	0.64	0.67	0.82	0.64	0.52	0.67	0.64
tvec + multi_nb (tuning)	0.74	0.67	0.66	0.65	0.69	0.66	0.52	0.73	0.66
hvec + multi_nb (tuning_2)	0.73	0.67	0.67	0.68	0.85	0.67	0.52	0.69	0.67
tvec + multi_nb (tuning_2)	0.72	0.67	0.69	0.66	0.69	0.67	0.52	0.73	0.67

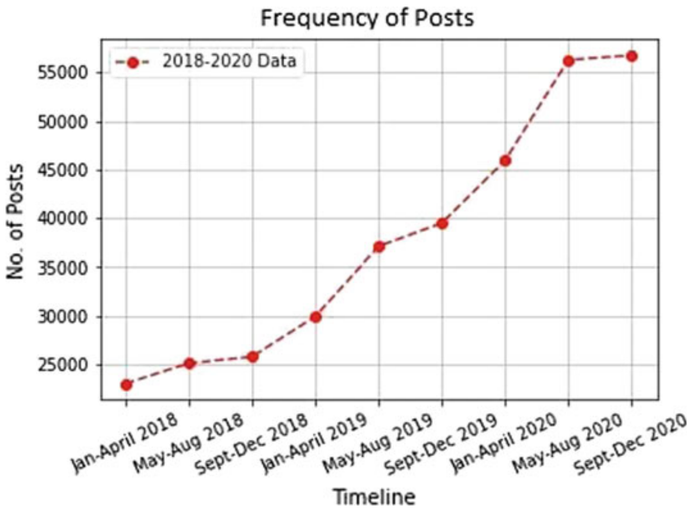


Fig. 2 Time series graph to see the change in the number of posts every year

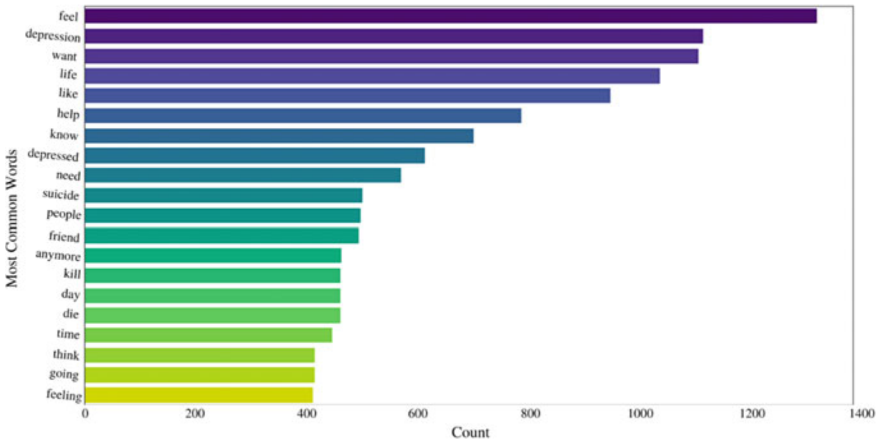


Fig. 3 Bar plot showcasing the frequency of most used words before covid, i.e., 2019

"covid," "coronavirus," "pandemic," and "covid-19" have emerged. This indicates that when Covid-19 hit the world, it also became a factor causing distress and affected the mental health of the people leading to a much more depressed and sad atmosphere. A recent study showed that social media consumption (72%) and posting (43%) increased during the pandemic and the content changed from positive to overwhelming (17%), information overload (15%), and stressful (15%) [16]. All this supports our time series analysis by telling that covid could have been a major factor in causing a significant increase in the number of posts during 2020.

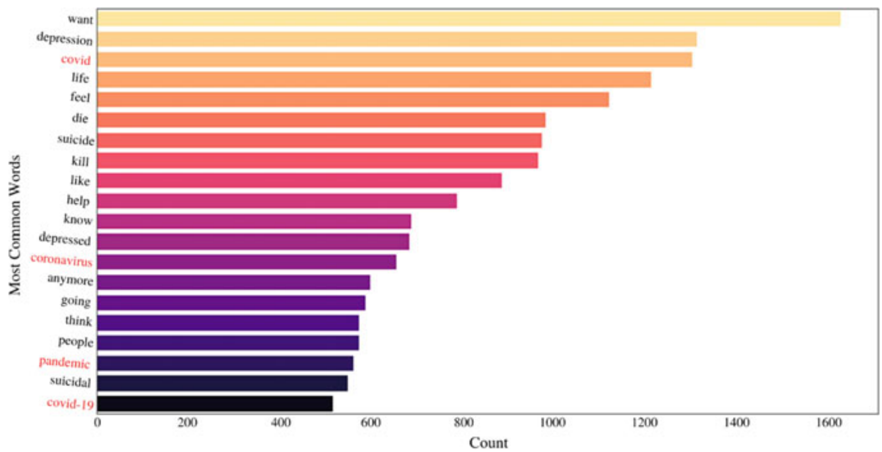


Fig. 4 Bar plot showcasing the frequency of most used words during covid, i.e., 2020

Fig. 5 Word cloud showcasing the most common words used **a** before Covid, i.e., 2019, **b** during Covid, i.e., 2020

4.1 Operationalizing The Classifier

To test our classifier on real-world examples, we took the book "Suicidal: Why We Kill Ourselves" [17] by Jesse Michael Bering who is an American psychologist. We applied the model to one of the excerpts from the book about a female patient. And the following were the predictions:

The model's predictions based on diary entries are depicted in this bar graph. The algorithm correctly recognized 70.6% of "pure" journal entries (those without letters, suicide notes, or poems) as "suicide." [18, 19] (Fig. 6).

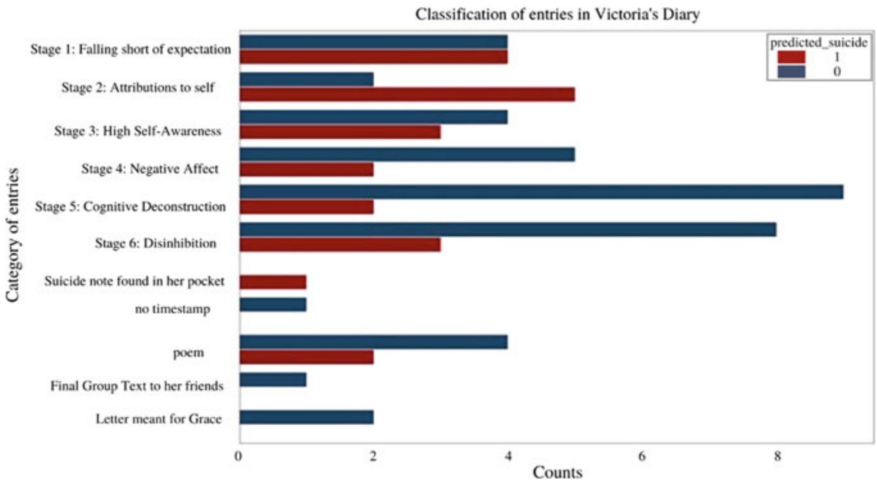


Fig. 6 Classification of entries in Victoria's diary

5 Conclusion

In this paper, we tried to research how Covid-19 mentally affected people all over the world. For the same, we examined the posts on the Reddit social media and particularly worked with two subreddits r/depression and r/SuicideWatch. Our findings showcased a significant increase in the number of posts during 2020. To support this statement, we generated a word cloud that displayed the most frequent words used before and during covid. The word cloud vividly indicated the presence of covid as the third most used word in the posts during 2020. There are undeniably many causes influencing the rise in the posts but one such cause could have been covid. One cannot simply disagree with how corona made people more distressed and suicidal, and our results are proof of that.

We first looked at the presence of overlapping terms in the dataset across both classes while analyzing the outcomes of the classifier model's application. Our application of TF-IDF did play a decent role in vectorizing those keywords and helping our Multinomial Naive Bayes classifier to work at its maximum potential.

6 Future Scope

6.1 Time Series Data Analysis

In our future work, we would like to experiment with the places affected such as how covid affected the people in America, Spain, London, and India. Ultimately, we would do a time series analysis to see how people's mind was affected for a specific country and whether it increased or decreased over time.

6.2 Chatbot

We believe that by incorporating the classifier model into a tailored chatbot, we can combine artificial intelligence with the empathy and competence of a psychologist to develop a strong mental health service that can reach millions of people worldwide. This would be useful for providing emotional support to the user by engaging with him or her and asking for frequent inputs related to the person's daily activities and feelings. This information can then be entered into our classification engine to learn more about the individual's mental state. Important contacts, such as family and friends, will be contacted if someone is having suicidal thoughts.

References

1. Depression. Who.int (2021) (online). Available https://www.who.int/news-room/fact-sheets/detail/depression. Accessed 25 Jul 2021
2. Yao H, Chen J, Xu Y (2021) Patients with mental health disorders in the COVID-19 epidemic
3. Sathish R, Manikandan R, Silvia Priscila S, Sara BV, Mahaveerakannan R (2020) A report on the impact of information technology and social media on covid-19. In: 2020 3rd international conference on intelligent sustainable systems (ICISS), pp 224–230. https://doi.org/10.1109/ICISS49785.2020.9316046
4. Khan MA, Uddin MF, Gupta N (2014) Seven V's of big data understanding big data to extract value. In: Proceedings of the 2014 zone 1 conference of the American society for engineering education, pp 1–5. https://doi.org/10.1109/ASEEZone1.2014.6820689
5. Prabhat A, Khullar V (2017) Sentiment classification on big data using Naïve Bayes and logistic regression. Int Conf Comput Commun Inform (ICCCI) 2017:1–5. https://doi.org/10.1109/ICCCI.2017.8117734
6. Solangi YA, Solangi ZA, Aarain S, Abro A, Mallah GA, Shah A (2018) Review on natural language processing (NLP) and its toolkits for opinion mining and sentiment analysis. In: 2018 IEEE 5th international conference on engineering technologies and applied sciences (ICETAS), pp 1–4. https://doi.org/10.1109/ICETAS.2018.8629198
7. Srividya M, Mohanavalli S, Bhalaji N (2018) Behavioral modeling for mental health using machine learning algorithms. J Med Syst 42:88. https://doi.org/10.1007/s10916-018-0934-5
8. Alonso SG, de la Torre-Díez I, Hamrioui S et al (2018) Data mining algorithms and techniques in mental health: a systematic review. J Med Syst 42:161. https://doi.org/10.1007/s10916-018-1018-2
9. Retrieved 18 May 2021, from https://www.reddit.com/r/depression, https://www.reddit.com/r/SuicideWatch
10. Khan N, Yaqoob I, Hashem I, Inayat Z, Mahmoud Ali W, Alam M et al (2014) Big data: survey, technologies, opportunities, and challenges. Sci World J 2014:1–18. https://doi.org/10.1155/2014/712826
11. Shankar VG, Devi B, Srivastava S (2019) Data speak: data extraction, aggregation, and classification using big data novel algorithm. In: Iyer B, Nalbalwar S, Pathak N (eds) Computing,

communication and signal processing. Advances in intelligent systems and computing, vol 810. Springer, Singapore. https://doi.org/10.1007/978-981-13-1513-8_1

12. Kowsher M, Tahabilder A, Hossain Sarker MM, Islam Sanjid MZ, Prottasha NJ (2020) Lemmatization algorithm development for Bangla natural language processing. In: 2020 Joint 9th international conference on informatics, electronics and vision (ICIEV) and 2020 4th international conference on imaging, vision and pattern recognition (icIVPR), pp 1–8. https://doi.org/10.1109/ICIEVicIVPR48672.2020.9306652
13. Ravindran S, Aghila G (2020) A data-independent reusable projection (DIRP) technique for dimension reduction in big data classification using k-nearest neighbor (k-NN). Natl Acad Sci Lett 43:13–21. https://doi.org/10.1007/s40009-018-0771-6
14. Surya PP, Seetha LV, Subbulakshmi B (2019) Analysis of user emotions and opinion usingmultinomial Naive Bayes classifier. In: 2019 3rd international conference on electronics, communication and aerospace technology (ICECA), pp 410–415.https://doi.org/10.1109/ICECA.2019.8822096
15. Chandrasekar P, Qian K (2016) The impact of data preprocessing on the performance of a Naive Bayes classifier. In: 2016 IEEE 40th annual computer software and applications conference (COMPSAC), pp 618–619. https://doi.org/10.1109/COMPSAC.2016.205
16. Digital Commerce 360 (2021) COVID-19 is changing how, why and how much we're using social media (online). Available at https://www.digitalcommerce360.com/2020/09/16/covid-19-is-changing-how-why-and-how-much-were-using-social-media/
17. Suicidal: Why We Kill Ourselves and Bering J (2021) Suicidal. Retrieved 18 May 2021, from https://www.goodreads.com/book/show/39105401-suicidal
18. Jain T et al (2020) Supervised machine learning approach for the prediction of breast cancer. In: 2020 international conference on system, computation, automation and networking (ICSCAN). IEEE
19. Yadav A et al (2021) Evaluation of machine learning algorithms for the detection of fake bank currency. In: 2021 11th international conference on cloud computing, data science and engineering (confluence). IEEE

State of Charge Estimation of the Lithium-Ion Battery Pack Based on Two Sigma-Point Kalman Filters

Nguyen Vinh Thuy, Nguyen Van Chi, Ngo Minh Duc, and Nguyen Hong Quang

Abstract Nowadays, the lithium-ion battery pack (LiB) is used as the main power supply for electric vehicles (EV). The remaining energy of LiB is the very important parameter determined continuously by estimating LiB's state of charge (SoC). SoC estimation is one of the main functions of the battery management systems (BMS). This article presents the use of two sigma-point Kalman filters (SPKF) to estimate accurately the SoC of the LiB based on the second-order model of the cell. The LiB's average SoC and the zero bias of the current measurement through the LiB are estimated by the first SPKF, while the second filter is applied to calculate the SoC differences between LiB's average SoC and the modules' SoC in the LiB. To improve the SoC accuracy of the LiB modules, a second-order RC equivalent circuit model (SECM) of the cell is used, and the influences of temperature, voltage hysteric, measurement errors, and zero bias of current measurement on the SoC estimation of the LiB are taken into account. To verify the method, the experimental test is conducted in the LiB with cells connected in parallels and series. The simulation and experimental results are analyzed to prove that the SoC estimation of the modules in the LiB is higher accuracy, and the LiB's average SoC errors are less than 1.5% at different temperatures ranging from − 5 to 45 °C. The calculation time consuming is shorter, and the calculation complex is reduced significantly.

Keywords Lithium-ion cell · Battery pack · Sigma-point Kalman filter · Current bias · SoC estimation · Second-order RC equivalent circuit model

1 Introduction

From the practical point of view, there are many advantages of the LiB such as higher energy density, less weight, high voltage out (about 3.7 V), safety, and fast charging/discharging rates comparing to the other kinds of battery [1, 2]. Today, the LiB is used more largely in the practice applications varying from electronics

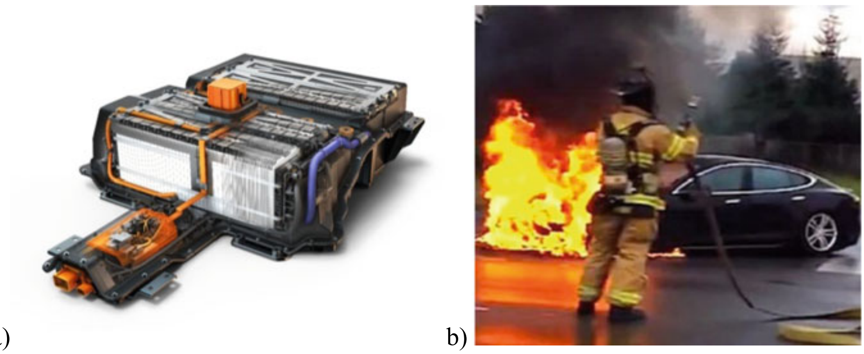
N. V. Thuy · N. V. Chi · N. M. Duc · N. H. Quang (✉)
Thai Nguyen University of Technology, Thai Nguyen, Vietnam
e-mail: quang.nguyenhong@tnut.edu.vn

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_37

devices like laptops, mobile phones, and small power home devices to the large power electrical vehicles. LiB operation is safe in the small power devices, but when it comes to the large power and high voltage applications, the LiB is easily unsafe because the LiB is formed by ten to thousands of cells connected in series-parallel in order to supply enough power (see Fig. 1a). The cell explosion caused by overcharge/over discharge could happen in the practice (see Fig. 1b). The control problem to ensure LiB operating stably, safety, and optimally is very important. This task is conducted by the BMS. The main functions of BMS are protecting the LiB, calculating the SoC, control charging/discharging, monitoring the health and safety of the LiB, etc. The SoC is the amount of energy remaining in the battery, and it is a significant input parameter of BMS and reflects the battery performance. The SoC cannot be measured directly but can be estimated by using the cell voltage, the current of the cell, the ambience temperature, etc. The BMS uses the accurate SoC estimation as an input not only to protect LiB, prevent the LiB from overdischarge/overcharge, and improve LiB life but also to conduct the control strategies to equalize energy level of cells in the LiB and to save energy [3].

Up until now, regarding SoC estimation methods for the LiB cell, there are many approaches being used. First, methods related to the model of cells in the LiB include the methods based on open circuit voltage, coulomb counting, and impedance method [4, 5]; in the works [6, 7], authors used the cell's first-order model to estimate the SoC; and in order to improve the accuracy of SoC estimation, the hysteristic, the aging process, and the change of cell internal resistance are considered in the SoC estimation [8]. The second, methods consider to the SoC estimation algorithms as the extended Kalman filters presented in the materials [9, 10], the particle filters used in the references [11, 12], the methods using learning algorithms presented in [13, 14], the methods based on fuzzy logic and nonlinear model used in the materials [15, 16].

The SoC estimation of the LiB is the complex issue, and many research works in the literature [17–20] related to this problem have been implemented in recent years. The SoC estimation is needed to be considered not only the complexity of the



a) b)

Fig. 1 **a** The electric car battery; **b** the battery fire of Tesla Model S

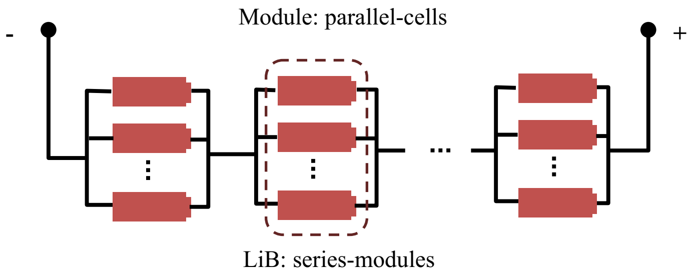


Fig. 2 Structure of the LiB

LiB's model and the complexity of the estimation algorithm but also the varying of parameters of the cell model and measurement noise.

There are many specific directions to solve the SoC estimation problem for LiB with distinct accuracy level, depending on the practical applications. To estimate the SoC of the cells in the LiB, the measurements of operating temperature, voltages, and current of the LiB are made available. In the practice, the measurements of current and voltage are affected by the noise; especially, the current is drifted by zero bias which is caused by the amplifier.

In this article, the SoC estimation method based on SPKF [21] for the LiB formed by a series of modules, each module consisted of some paralleled cells as shown in Fig. 2, is presented. This SoC estimation takes into account the noises and the zero bias of the measurements of current and voltage.

In this work, to describe the cell dynamic we use the SECM. The noises and the zero bias of the measurements of current and voltage are considered in this model. The cell dynamic is reflected more exactly in the operation condition with charge and discharge magnitude varying suddenly by using the second-order RC equivalent circuit model. The SoC estimation algorithm based on two filters is used to estimate SoC for all modules in the LiB by summing LiB's average SoC estimated by the first SPKF filter and SoC difference of modules calculated by the second SPKF filter.

The remainder of paper is organized as follows: In part 2, the dynamic second-order RC equivalent circuit model of the LiB is presented. In part 3, the SoC estimation of LiB using two filters based on the SPKF is given. Some simulations and experimental results are shown in the next part. In the last part, discussions for the paper are mentioned.

2 The Dynamic Model of the Cell and the LiB

2.1 The Second-Order RC Equivalent Circuit Model of Cell

To describe more accuracy of the dynamic of the cell, especially for the application using LiB with charge/discharge amplitude varying suddenly, the SECM model is used based on our previous work [6].

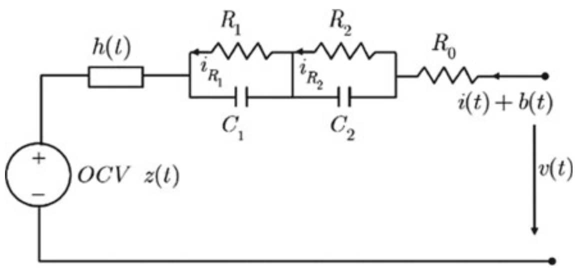
The SECM of the cell is depicted in Fig. 3. The notations in the model are as follows: two currents of two dynamic branches of the cell i_{R_1}, i_{R_2}; the hysteresis $h(t)$; SoC $z(t)$; zero bias of current of the cell $b(t)$; and$i(t)$ and $v(t)$ denote the current and the voltage of the cell, respectively. They are available by measuring and affected by noises.

In discrete-time domain, define $\underline{x}_{k+1}, \underline{u}_k, y_k$ to be the state vector, input and output vectors at the sample time k, $k = 0, 1, 2, \ldots, \infty$, respectively. They are written in Eq. (1).

$$\underline{x}_{k+1} = \begin{bmatrix} i_{R_1,k+1} \\ i_{R_2,k+1} \\ h_{k+1} \\ z_{k+1} \\ b_{k+1} \end{bmatrix}, \underline{u}_k = \begin{bmatrix} i_k \\ \mathrm{sgn}(i_k) \end{bmatrix}, y_k = v_k \tag{1}$$

The state model of the cell is written as shown in Eq. (2):

Fig. 3 The SECM of the cell



$$\begin{bmatrix} i_{R_1,k+1} \\ i_{R_2,k+1} \\ h_{k+1} \\ z_{k+1} \\ b_{k+1} \end{bmatrix} = \begin{bmatrix} A_{R_1C_1} & 0 & 0 & 0 & 0 \\ 0 & A_{R_2C_2} & 0 & 0 & 0 \\ 0 & 0 & A_h & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} i_{R_1,k} \\ i_{R_2,k} \\ h_k \\ z_k \\ b_k \end{bmatrix} + \begin{bmatrix} \left(1 - A_{R_1C_1}\right) & 0 \\ \left(1 - A_{R_2C_2}\right) & 0 \\ 0 & (1 - A_h) \\ \frac{-\eta_k \Delta t}{Q} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} i_k \\ \mathrm{sgn}(i_k) \end{bmatrix} + \begin{bmatrix} w_{i_{R1,k}} \\ w_{i_{R1,k}} \\ 0 \\ 0 \\ w_{b,k} \end{bmatrix} \tag{2}$$

The matrices in Eq. (2) are defined as:

$$\begin{aligned} & A_{R_1C_1} = \exp\left(\frac{-\Delta t}{R_1C_1}\right), A_{R_2C_2} = \exp\left(\frac{-\Delta t}{R_2C_2}\right), \\ & B_{R_1C_1} = 1 - A_{R_1C_1}, B_{R_2C_2} = 1 - A_{R_2C_2}; A_h = \exp\left(-\left|\frac{\eta_k i_k \gamma_k \Delta t}{Q}\right|\right); B_h = 1 - A_h \\ & \underline{w} = \left[\, w_{i_{R1,k}} \; w_{i_{R1,k}} \; 0 \; 0 \; w_{b,k} \,\right]^T \end{aligned} \tag{3}$$

in which $\underline{w}$ is the disturbance vector of the model formed by the current noises of two RC dynamic branches of cell $w_{i_{R_1,k}}$, $w_{i_{R_2,k}}$ and current bias noise of cell $w_{b,k}$. Two parameters depending on the ambient temperature of cell η_k and γ_k are the columbic efficiency.

The output equation of the SECM model is written as Eq. (4), and this equation describes the relationship between v_k and i_k, SoC, h_k, the currents of two RC branches $i_{R_1,k}$, $i_{R_2,k}$, and the voltage noise of cell ζ_k.

$$v_k = \mathrm{OCV}(z_k) + Mh_k - R_1 i_{R_1,k} - R_2 i_{R_2,k} - R_0 i_k + \zeta_k \tag{4}$$

Based on Eqs. (2), (3), and (4), the state space model of cell is:

$$\begin{cases} \underline{x}_{k+1} = A_k \underline{x}_k + B_k \underline{u}_k + \underline{w}_k \\ y_k = OCV(z_k) + C_k \underline{x}_k + D_k \underline{u}_k + \zeta_k \end{cases} \tag{5}$$

In the state space model (5), the matrices are formed as follows:

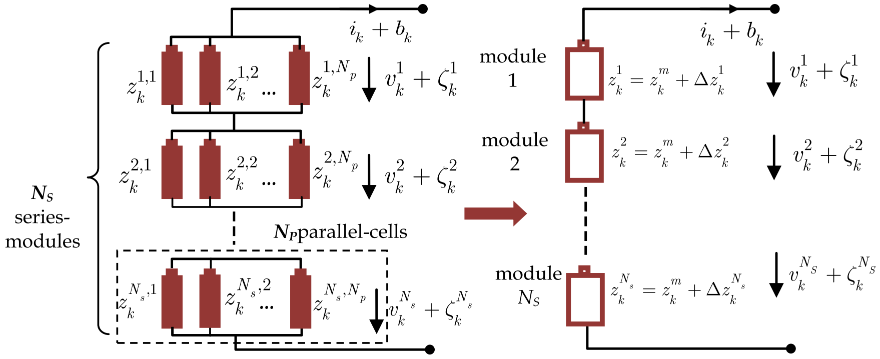


Fig. 4 The model of the lithium-ion battery pack (LiB)

$$A_k = \begin{bmatrix} A_{R_1C_1} & 0 & 0 & 0 & 0 \\ 0 & A_{R_2C_2} & 0 & 0 & 0 \\ 0 & 0 & A_h & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad B_k = \begin{bmatrix} \left(1 - A_{R_1C_1}\right) & 0 \\ \left(1 - A_{R_2C_2}\right) & 0 \\ 0 & (1 - A_h) \\ \frac{-\eta_k \Delta t}{Q} & 0 \\ 0 & 0 \end{bmatrix} \tag{6}$$
$$C_k = \left[-R_1 \ -R_2 \ M \ 0 \ 0\right], \quad D_k = \left[-R_0 \ 0\right]$$

2.2 The Dynamic Model of the LiB

The LiB is formed by some modules connected in series as described in Fig. 4. In this LiB structure, N_s is number of modules, and N_p is number of paralleled cells in the modules. The symbols of the quantities are described in Fig. 4 also.

For the paralleled cells, after a certain period of time cells's SoC will balance by itself. So the model of the LiB is transformed into the string with many modules connected in series, as plotted in Fig. 4. Demonstration of self-balance voltage and SoC of cells in one module is plotted in Figs. 5 and 6. Suppose that the initial SoCs of 6 cells vary in the range of 65–90%. The internal resistances of cells R_0 vary in the range of 1.0–1.4 mΩ. After 50 s, all six cells have the same voltage and the same SoC level.

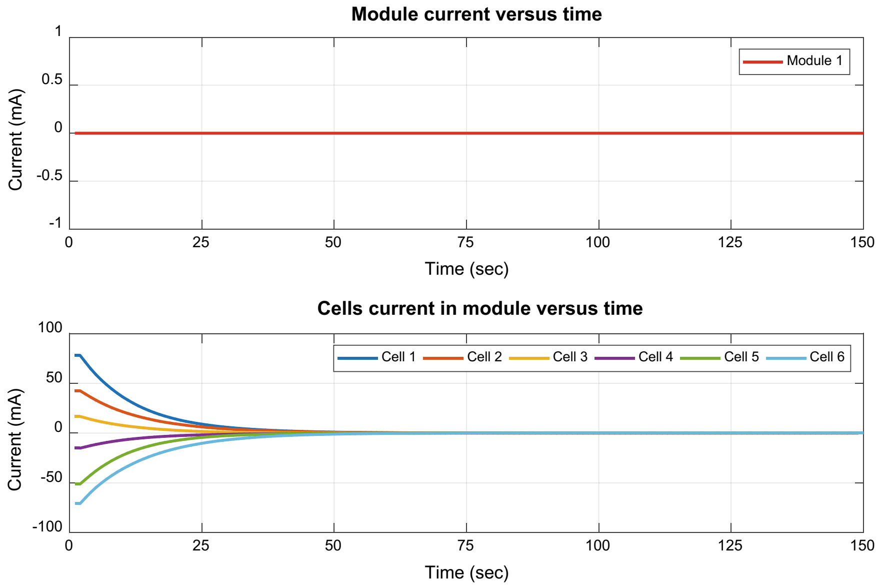


Fig. 5 **a** Current through the module, **b** current through the cells of the module

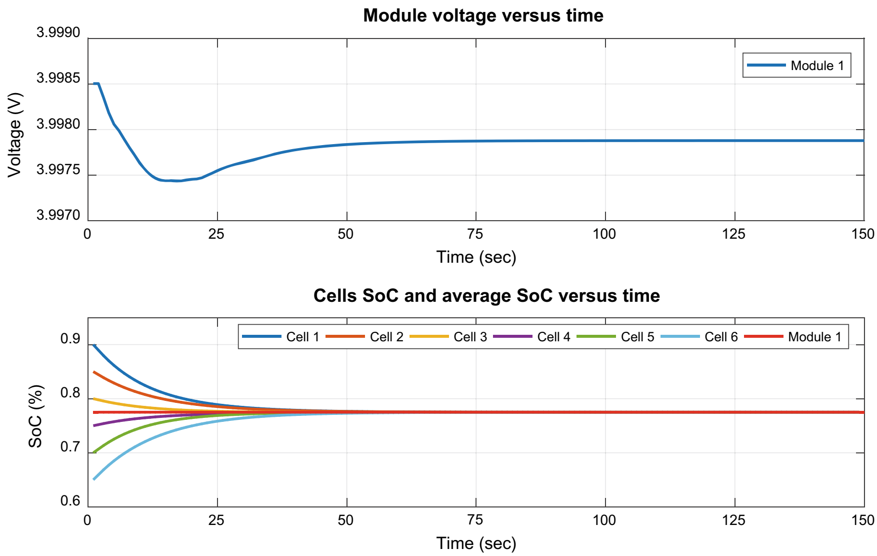


Fig. 6 **a** The voltage of the module, **b** SoC of the cells of the module

3 SoC Estimation of the LiB Using Two SPKF

The SoC estimation algorithm for the LiB using two SPKF is described as follows:

- **Step 1**: estimate the SoC average of the LiB using the first SPKF
- **Step 2**: estimate the SoC differences between the SoC average of the LiB and SoCs of the modules
- **Step 3**: The SoC of modules in the LiB is calculated by summing the SoC average (*estimated in the step 1*) and the SoC differences (*estimated in the step 2*)

To implement step 1, consider the LiB to be an equivalent cell that has the second-order RC equivalent circuit model as presented in Eq. (5). State variables need to be estimated at every sample times k are the two currents $i_{R_1}(k)$, $i_{R_2}(k)$ of two dynamic branches RC of the equivalent cell, the voltage hysteretic $h(k)$, the module's SoC $z(k)$, and the current's zero bias of $b(k)$.

The model input is the current of the LiB, it is affected by zero bias and measurement noise, and the output of the model is the sum of the voltages of modules in the LiB. We use the state vector formed as following equation:

$$\underline{x} = \left[i_{R_{1,k}} \ i_{R_2} \ h_k \ z_k \ b_k \right]^T \tag{7}$$

Define the notes $\sigma_{\tilde{x}}$, σ_w, and σ_ζ as covariance matrices of state estimation errors, systems noises, and voltage noise, respectively. The SoC estimation algorithm for the LiB is presented as following part.

State of Charge Estimation Algorithm for the LiB

Initialize the parameters of LiB
Initialize $\text{SoC}_0 \in R^{(N_s \times N_p)}$, $R_0 \in R^{(N_s \times N_p)}$
Initialize $\sigma_{\tilde{x}}$, σ_w, and σ_ζ
Calculate $\overline{z}_0^{(i)}$, $\overline{R}_0^i$, and Q^i
For$k = 1$ **to** ∞ **do**

Measure i_k, v_k, T_k
Do step 1: Estimate $\overline{z}_k$ by the first filter
For $i = 1$ to N_s do

Do step 2: Estimate $\Delta\hat{z}_k^{(i)}$ of module ith by using the second filter

End
Do step 3: Calculate $\hat{z}^{(i)} = \hat{z}_k + \Delta\hat{z}_k^{(i)}$, $i = 1, 2..., N_s$

End

4 Experimental SoC estimation results

In this study, the SoC estimation algorithms are coded in MATLAB software. The LiB is formed by seven serial modules; each module consists of six SAMSUNG ICR18650-22P paralleled cells. The technical parameters of cell are shown in the Table 1. The experimental system is depicted in Fig. 7. The SoC estimation is conducted with the LiB with a scenario of charge/discharge varying continuously in 1 h and current amplitude changing suddenly, and the maximum of discharge/charge current amplitude are 10 A and 3 A, respectively. The scenarios of charge/discharge in this test simulate the charge/discharge situations of LiB used in the EV in the practice. Suppose that the zero bias of the current $b(k)$ of LiB varies in the range of 0.1–0.5 A.

The current of the LiB in the charge/discharge scenarios is plotted in Fig. 8. The voltages of the modules in the LiB are shown in Fig. 9, and SoC varying by the time of the modules is described in Fig. 10. Figure 11 is the varying of the voltages and

Table 1 The technical parameters of LiB SAMSUNG ICR18650-22P

Item	Specifications
Model	ICR18650-22P
Nominal capacity	2150 mAh
Minimum capacity	2050 mAh (0.2 °C discharge, 2.75 V discharge)
Charging voltage	4.2 $\pm$ 0.05 V
Nominal voltage	3.62 V (1 °C discharge)
Charging current	1075 mA
Max. charge current	2150 mA
Max. discharge current	10 A (continuous discharge)
Discharge cut-off voltage	2.75 V
Operating temperature	Charge: − 10 to 50 °C; discharge: − 20 to 70 °C

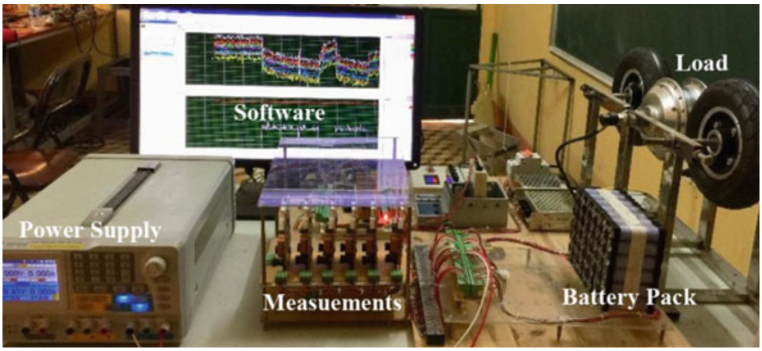


Fig. 7 Setup of the experimental system (A4 building, RIAT, TNUT)

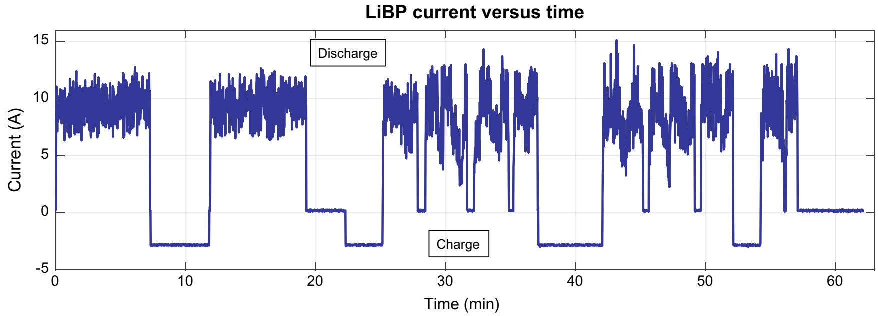


Fig. 8 The scenario of the discharge/charge current of the LiB

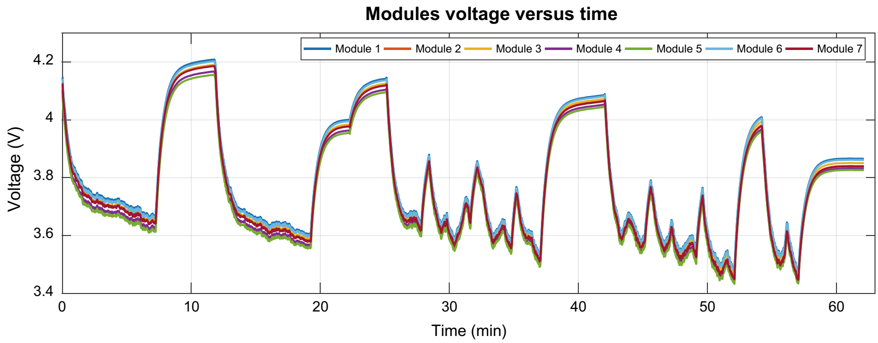


Fig. 9 The voltages of the modules in the LiB

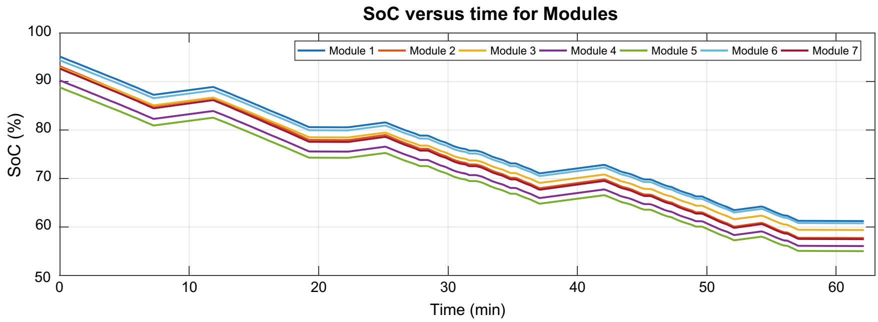


Fig. 10 The SoC varying by the time of the modules

currents of the cells in the module. The current of the 6 cells has the same rule as the current of the LiB, and the current of the LiB is equal to the total current flowing through all the cells.

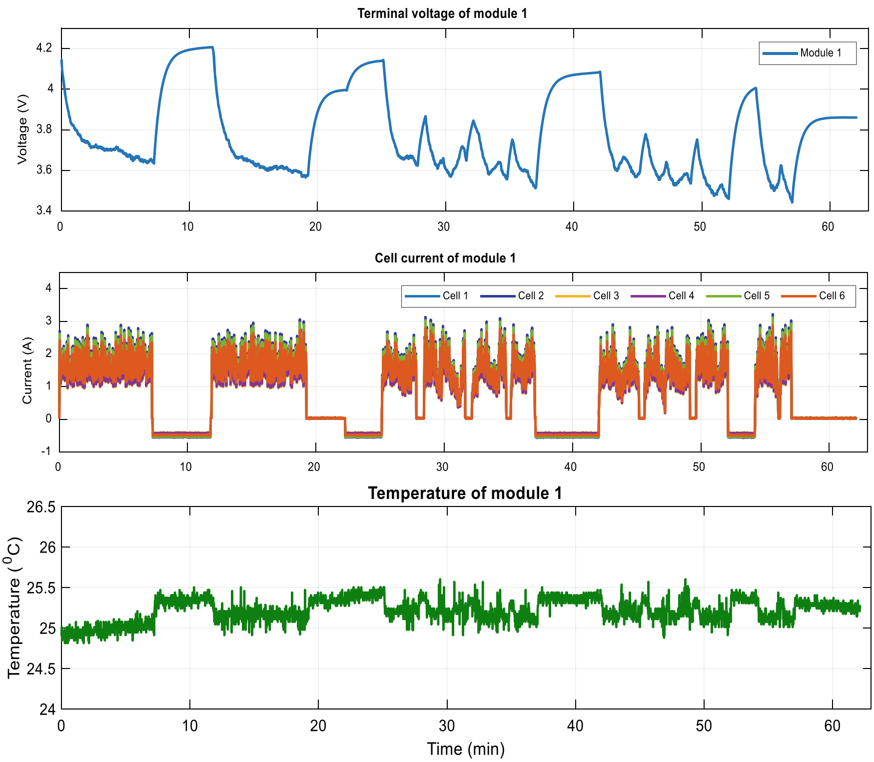


Fig. 11 The varying of the voltage and the currents of the cells in the module

Table 2 Model's dynamic parameters of the cell

T °C	− 5 °C	5 °C	15 °C	25 °C	35 °C	45 °C
η_k	1.0869	0.9803	1.0220	1.0183	1.0542	1.0399
Q (Ah)	2.1596	2.1877	2.1943	2.1507	2.1515	2.1523
γ	250.000	78.4915	63.6762	2.0748	170.6407	151.3064
M (V)	0.0347	0.0257	0.0188	0.0177	0.0201	0.0185
M_0 (V)	0.0072	0.0049	0.0048	0.0018	0.0036	0.0024
R_0 (Ω)	0.0013	0.0013	0.0012	0.0012	0.0012	0.0011
R_1 (Ω)	0.0204	0.0203	0.0201	0.0019	0.0019	0.0019
R_2 (Ω)	0.0494	0.0376	0.0288	0.0443	0.0136	0.0134
R_1C_1 (s)	0.6124	1.7555	0.3227	1.4881	0.2997	0.4630
R_2C_2 (s)	3.9035	7.5994	8.1118	36.8543	5.1840	6.5319

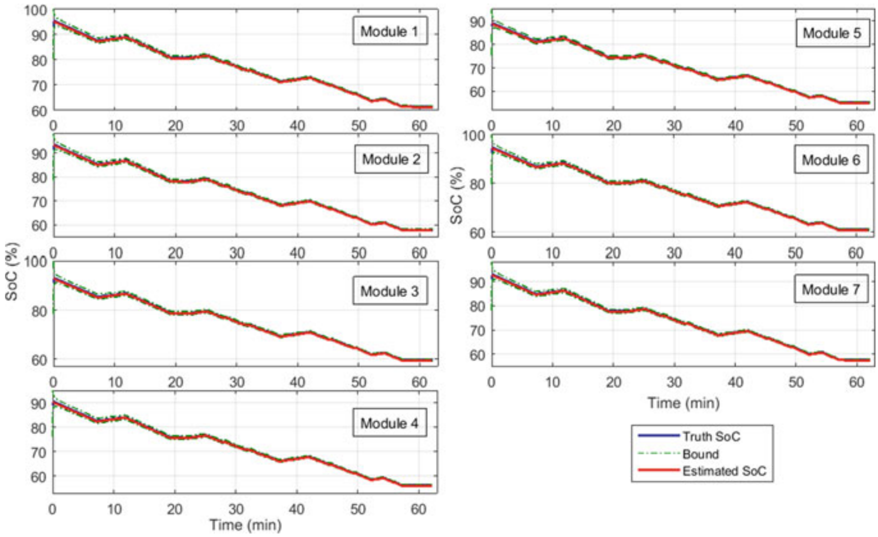


Fig. 12 The results of SoC estimation for the modules in the LiB

The dynamic parameters of the cell are given in Table 2. The initial parameters of the LiB determined at $T = 25\,^{\circ}\mathrm{C}$ are $\mathrm{SoC}_0 \in R^{(N_s \times N_p)}$, $R_0 \in R^{(N_s \times N_p)}$, and $Q_0 \in R^{(N_s \times N_p)}$. The covariance matrices are:

$$\sigma^{+}_{\tilde{x},0} = \begin{bmatrix} 0.001 & 0 & 0 & 0 & 0 \\ 0 & 0.0001 & 0 & 0 & 0 \\ 0 & 0 & 0.01 & 0 & 0 \\ 0 & 0 & 0 & 0.01 & 0 \\ 0 & 0 & 0 & 0 & 0.01 \end{bmatrix}, \sigma_{\tilde{w}} = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix}, \sigma_{\tilde{v}} = 0.0001$$

The estimated average SoC for each module in the LiB is depicted in Fig. 12, and the SoC estimation error of the modules is shown in Fig. 13. The estimated SoC shows that the SoC estimation of the modules when considering the zero bias of the current of the LiB has been tracked to the actual SoC average of the LiB with the estimated SoC error in the test is quite small, about 0.28% for each module.

The estimated current's zero bias of LiB is plotted in Fig. 14. The real values of zero bias are set as 0.1 A, 0.3 A, and 0.5 A, respectively. After a period of time $t =$ 2 min, the estimated zero bias tracks up to the real value and it is distributed around the real value of zero bias. The average value of the estimated zero bias in the test is 0.102 A, 0.306 A, and 0.484 A, respectively, with the errors of 2.0%, 2.0%, and 3.2%. These are quite small errors, so this test shows that the estimation of zero bias is suitable for practical applications.

The SoC estimation results and estimated errors for module 1 as shown in Fig. 15 when the zero bias of the LiB current varying from 0.0 A to 0.5 A at the operation

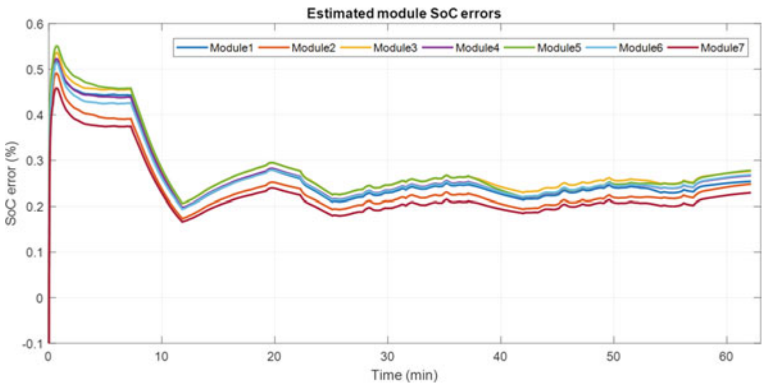


Fig. 13 The results of SoC estimation errors of the modules in the LiB

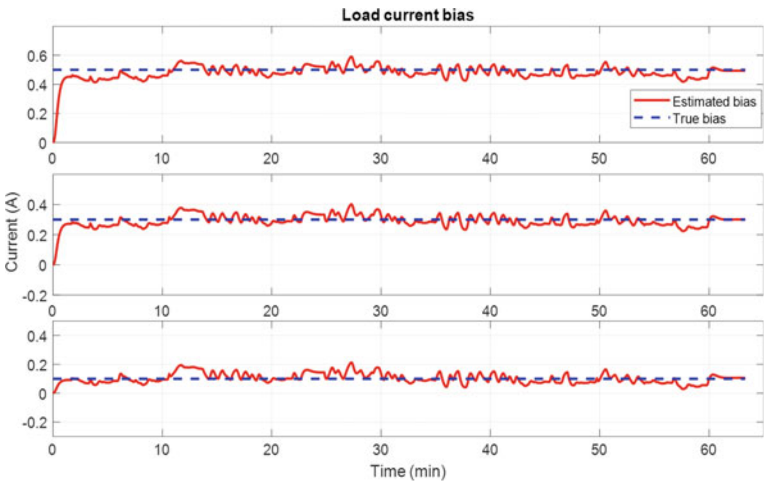


Fig. 14 The estimated result of the zero bias of the current of LiB

temperature $T = 25$ °C. The comparison of the SoC estimation errors of the LiB over the experimental period is illustrated in Table 3.

Figure 16 shows the SoC estimation error of the modules in the LiB according to the operating temperature at $T = [-5$ °C $\div$ 45 °C] with the zero bias = 0.3 A. From the above of the SoC estimation results, it shows that the estimation errors of the modules in LiB are small; in another word, the estimation of SoC has high accuracy. In the temperature ranging from – 5 to 45 °C, the SoC estimation errors of the modules are less than 1%. When the working temperature of the LiB is decreased to – 5 °C, the SoC estimation errors of the modules have an increasing tendency, but the largest error value is less than 2.3% and the average error value is about 1.5%.

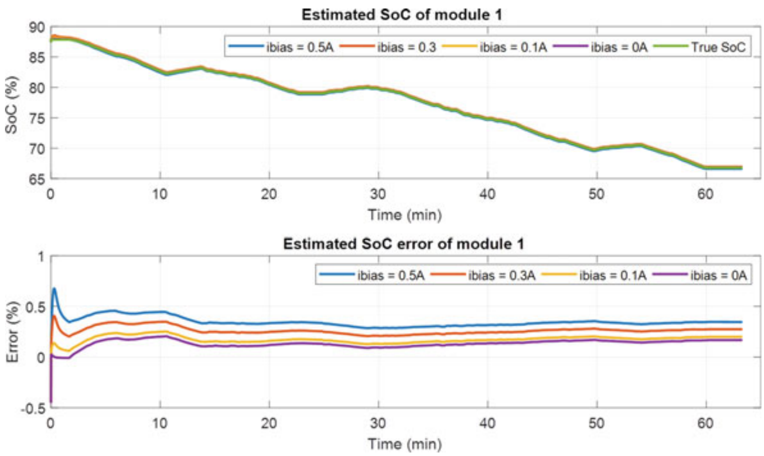


Fig. 15 The SoC estimation results and estimated errors for the module 1

Table 3 The comparison of the SoC estimation errors with respect to values of the zero bias of current of the LiB

I bias (A)	SoC error (%), $T = 25$ °C						
	Module 1	Module 2	Module 3	Module 4	Module 5	Module 6	Module 7
$b_k = 0.0$	0.12	0.10	0.12	0.11	0.14	0.11	0.12
$b_k = 0.1$	0.21	0.20	0.22	0.21	0.24	0.20	0.21
$b_k = 0.3$	0.39	0.38	0.40	0.40	0.43	0.38	0.40
$b_k = 0.5$	0.57	0.56	0.59	0.60	0.63	0.56	0.59

This is a very important SoC estimation error for the SoC estimation problem for the LiB in EV applications.

5 Conclusion

This paper presented a method to improve the SoC estimate accuracy for a LiB including many cells that are connected in series and parallel. This study uses two filters based on the SPKF algorithm to design the SoC estimation method for the LiB when taking into account the effect of temperature, measurement noise, and zero bias of current of the LiB. The dynamic model of cell in the LiB is described by the SECM to reflect more accurately the nonlinear characteristics of the LiB. The SoC estimation algorithm is applied experimentally to LiB, this LiB is formed by ICR 18650-22P SAMSUNG cells in 7 serial modules, and each module consists of 6 parallel cells. The SoC estimation results for the LiB under with the temperature changing from – 5 to 45 °C show that the errors of SoC estimations for modules in

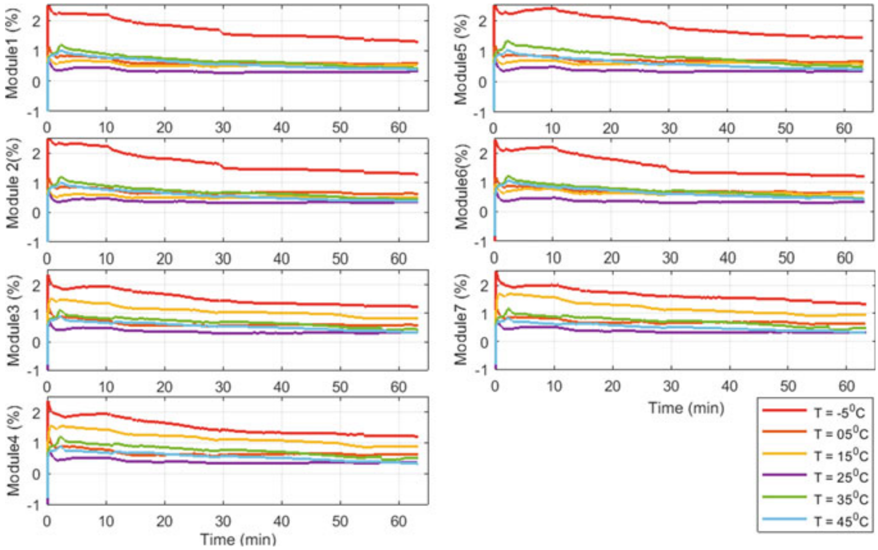


Fig. 16 The results of SoC estimation error of the modules in the LiB with respect to operating temperatures

the LiB are quite small, and the accuracy of SoC estimation has been significantly improved compared to other methods. The estimation method in this study can be applied to the SoC estimation problem for the LiB with a large number of cells. The calculation complex is reduced. This result is significant when the LiB for EV today is made by thousands of cells. Our future work focuses on improving the accuracy of the current zero bias estimation of the LiB.

Acknowledgements This research was financially supported in T2020-B12 project by Thai Nguyen University of Technology—TNUT.

References

1. Zhou Y, Li X (2015) Overview of lithium-ion battery SOC estimation. In: IEEE international conference on information and automation, pp 2454–2459
2. Horiba T (2014) Lithium-ion battery systems. Proc IEEE 102:939–950
3. Xiong R, Hongwen H (2014) Cell state-of-charge estimation for the multi-cell series-connected battery pack with model bias correction approach. Energy Proc 61:172–175
4. Ng KS, Moo CS, Chen YP, Hsieh YC (2009) Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries. Appl Energy 86:1506–1511
5. Zheng Y, Ouyang M, Han X, Lu L, Li J (2018) Investigating the error sources of the online state of charge estimation methods for lithium-ion batteries in electric vehicles. J Power Sour 377:161–188

6. Thuy NV, Chi NV (2020) State of charge estimation for lithium-ion battery using sigma-point Kalman filters based on the second order equivalent circuit model. Advances in engineering research and application. ICERA 2019. Lecture notes in networks and systems. Springer, Cham, p 104
7. HuX LiS, ZhangC PH (2012) A comparative study of equivalent circuit models for li-ion batteries. J Power Sour 198:359–367
8. Yazdanpour M, Taheri P, Mansouri A, Schweitzer B (2016) A circuit-based approach for electro-thermal modeling of lithium-ion batteries. In: 32nd thermal measurement, modeling & management symposium (SEMI-THERM), pp 113–127
9. Xu Y, Hu M, Zhou A, Li Y, Li S, Fu C, Gong C (2020) State of charge estimation for lithium-ion batteries based on adaptive dual Kalman filter. Appl Math Model 77:1255–1272
10. Ciortea F, Rusu C, Nemes M, Gatea C (2017) Extended Kalman filter for state-of-charge estimation in electric vehicles battery packs. In: International conference on optimization of electrical and electronic equipment (OPTIM) & 2017 international Aegean conference on electrical machines and power electronics (ACEMP), pp 611–616
11. Zhang M, Wang K, Zhou YT (2020) Online state of charge estimation of lithium-ion cells using particle filter-based hybrid filtering approach. Hindawi. https://doi.org/10.1155/2020/8231243
12. Tran NT, Vilathgamuwa M, Li Y, Farrell T, Choi SS, Teague J (2017) State of charge estimation of lithium ion batteries using an extended single particle model and sigma-point Kalman filter. IEEE Southern Power Electron Conf (SPEC). https://doi.org/10.1109/SPEC.2017.8333564
13. Tong S, Lacap JH, Park JW (2016) Battery state of charge estimation using a load- classifying neural network. J Energy Storage 7:236–243
14. Sassi HB, Errahimi F, Es-Sbai N, Alaoui C (2019) Comparative study of ANN/KF for on-board SoC estimation for vehicular applications. J Energy Storage 27
15. Jiani D, Zhitao L, Youyi W, Changyun W(2014) A fuzzy logic-based model for Li-ion battery with SoC and temperature effect. In: 11th IEEE international conference on control and automation (ICCA), pp 1333–1338
16. Wang B, Liu Z, Li SE, Moura SJ, Peng H (2017) State-of-charge estimation for lithium-ion batteries based on a nonlinear fractional model. IEEE Trans Control Syst Technol 25:3–11
17. Safi J, Beeney M, Kehs M, Anstrom J, Brennan S, Fathy H (2014) Improving SOC accuracy using collective estimation for lithium ion battery cells in series. In: American control conference, pp 254–259
18. He F, Shen WX, Kapoor A, Honnery D, Dayawansa D (2016) H infinity observer based state of charge estimation for battery packs in electric vehicles. In: IEEE 11th conference on industrial electronics and applications (ICIEA), pp 694–699
19. Mawonou KSR, Eddahech A, Dumur D, Godoy E, Beauvois D, Mensler M (2018) Li-ion battery pack SoC estimation for electric vehicles. IEEE, pp 4968–4973
20. Tran NT, Khan AB, Nguyen TT, Kim DW, Choi WJ (2018). SOC estimation of multiple lithium-ion battery cells in a module using a nonlinear state observer and online parameter estimation. Energies 11(7):1620. https://doi.org/10.3390/en11071620
21. Van CN (2019) State estimation based on sigma point Kalman filter for suspension system in presence of road excitation influenced by velocity of the car. J Control Sci Eng 2019:1–16

Image Processing Technique in Measuring Underwater Target's Properties

Intisar F. H. Al-Shimiry, Ali A. D. Al-Zuky, and Fatin E. M. Al-Obaidi

Abstract Due to light refraction effect, the need to design a proper technique for underwater scene is a critical task. The problem began to be appeared at deeper levels of water and hence, one cannot distinguish nor observe the properties for underwater objects. The properties may include dimensions, color, texture, etc. So, the current paper involves such situation through designing a system prepared for this purpose with the use of two types of water; an ordinary type with an addition of alum and distilled water. The presented work aims to measure object's properties of three different objects immersed inside such system. Results show a complete match between the actual and estimated values for the relationship between the two used laser spots distance-water depth variation for the two used sources of water. In addition to that, a special convergence appeared clearly upon the actual and estimated values in detecting underwater object's properties especially at water depth that exceed one meter.

Keywords Underwater image · Distance-water depth variation · Scaling factor · Distance between two laser spots

1 Introduction

For many applications including geological, archaeological, physical, biological, and industrial issues, the exploration of underwater environments is an important task [1]. Nevertheless, this exploration faces some difficulties related to some constraints. Such constraints may include view-disturbing noises, light attenuation, and refraction effects. In water, the intensity of light decreases with the distance from the object

I. F. H. Al-Shimiry · A. A. D. Al-Zuky · F. E. M. Al-Obaidi (✉)
Department of Physics, College of Science, Mustansiriyah University, Baghdad, Iraq
e-mail: sci.phy.fam@uomustansiriyah.edu.iq

I. F. H. Al-Shimiry
e-mail: entsarfalah@uomustansiriyah.edu.iq

A. A. D. Al-Zuky
e-mail: prof.alialzuky@uomustansiriyah.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_38

and such effect depends on light's wavelength and hence, the red light shall decrease in water easier than the blue one. Some problems related to the suspended materials immersed in water and constraint a noisy source like fishes, bubbles, etc. The third problem related to the effects of light refraction caused by two different media of different refractive indices [2]. Because of light's attenuation, underwater scenes suffer from haze and as a result, a poorly contrasted images resulted. Such limitation disturbs the visibility distance to 20 and 5 m or less for clear and turbid water, respectively [3].

The originality of this work focusses upon a comparison study between the actual and estimated values for the statistical properties of underwater targets under different states of water depths.

Many researches focused their attention toward detecting and measuring underwater object's properties. The authors in [4] studied triangulation, telemetry, and interferometry to measure distance, surface with their applications, properties and then subtend each of them. Mrovlje J. and Vrančić D. calculated object's position using geometrical derivations. Results show a relatively accurate measurements in calculated distance to the subject under study [5]. Al-Obaidi in [6] studied the theoretical and experimental properties for different water types by utilizing digital image processing and found a matching in behavior between certain water types in their bands. Related to previous work, Al-Shimiry et al. in [7] found an accuracy in describing object's properties (scaling factor, magnification, and object's length) for submersible and floating objects. In addition to the latter work, the authors in [8] predict an algorithm for measuring water depth estimation with the aid of image processing. They introduced four geometric models with progressive curves in distance-water depth variations. Mohamed H et al. proposed three empirical models for bathymetry calculations, and assessed their performance with Landsat 8 and spot 6 satellite images by using bagging (BAG) algorithm. Such technique seems at last the most accurate and preferable algorithm than other used algorithms in bathymetric estimation [9]. Al_Mawlawi H. M. J. utilized two laser pointers to specify object's dimensions in addition to determine its distance in an open air environment. She found a good matching with high accuracy between the experimental and calculated results [10]. Chybicki A. introduced some models of Satellite Derived Bathymetry (SDB) for turbid water of South Baltic Sea. Their results analyzed in terms of spatial distribution, overall quality, and depth error estimation. Such models can be applied successfully for the South Baltic at depths equal to 12–18 m with an error up to 10–20% of the real depth [11]. Dietrich J. T. used Structure-from Motion (SfM) and estimated series of refraction correction equations for each point/camera combination in SfM point cloud. The technique gave an accuracy of ~ 0.02% for flying height with ~ 0.1% precisions [12].

2 The Proposed System

The designed system consists of a cylindrical plastic tank of (120 × 100) cm. Its interior surface is painted by black pigment. Two laser pointers of red colors had been used, here their wavelengths are 6100 A°. Such designed system is prepared to measure target's properties for three different objects that immersed at tank's bottom. These objects include a square object of 15 cm in length, triangle (12 × 21, 24) cm, and (20 × 10) cm for a rectangle object as shown in Fig. 1. Images had been captured for each water depth ranging from 0 to 100 cm with an increment equal to (10) cm for each step. The used camera is a Samsung's camera phone with super speed dual pixel 12 MP AF sensor with 77° field of view and dual apertures f/1.5 mode, f/2.4 mode.

The presented work consists of two stages, the first part by using an ordinary water with the addition of alum which is a chemical compound consisting of two simple salts, that is hydrated potassium sulfate and aluminum sulfate, called potassium salts, as they dissolve in water and filters the water and finally increases its purity. The second section included the using of 1000 L of distilled water. The whole technique for measuring the properties of underwater targets used the click properties which can detect the targets found in the scene as t_1, t_2, t_3, and t_4 for laser two spots, square, triangle, and rectangle objects, respectively. A pre-processing operation had been executed, here by adopting histogram equalization to the captured image to enhance and distribute image gray levels all over the range and hence, the next equations can be applied to get object's properties.

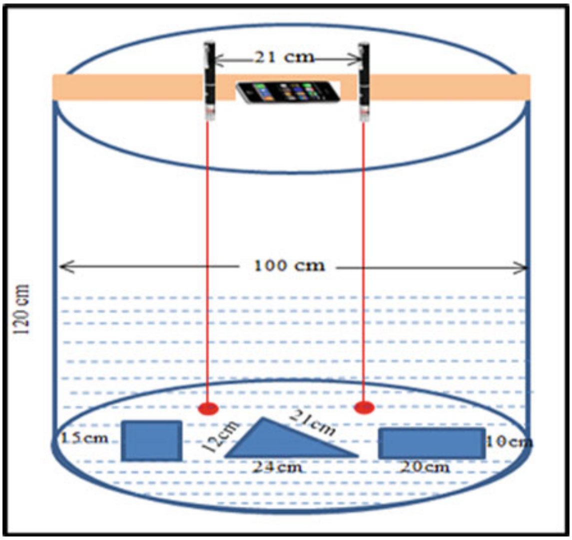


Fig. 1 Proposed system for measuring underwater object's properties

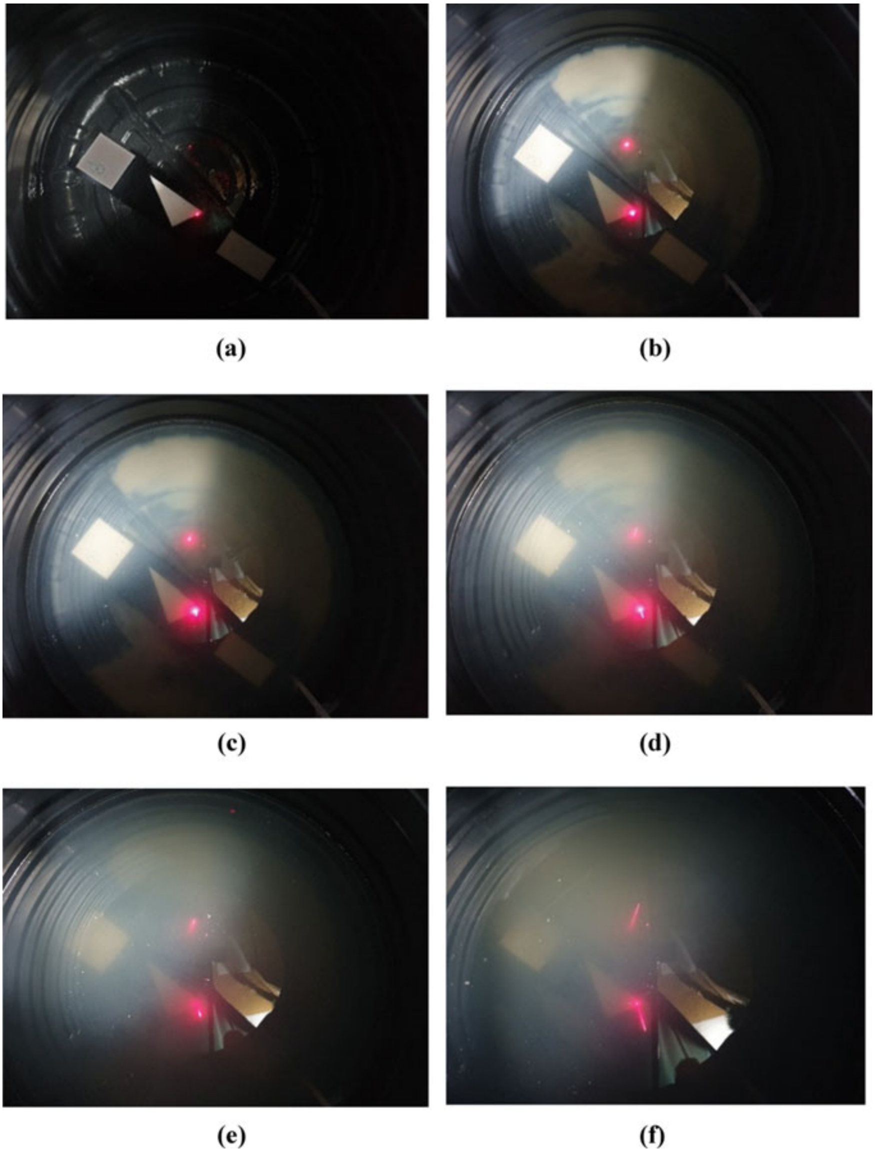
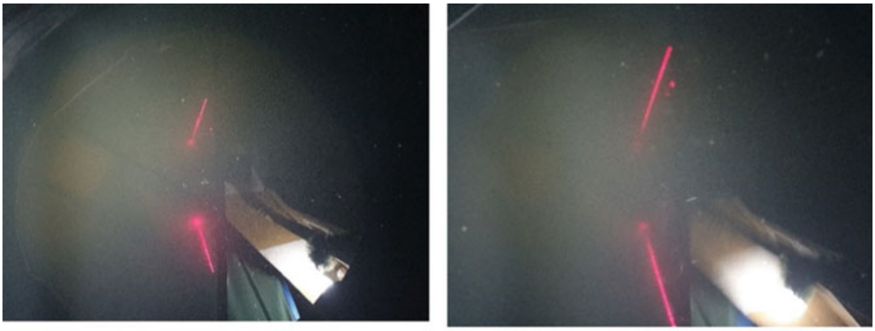


Fig. 2 Captured images by using potassium salts in water for water depth equal to: **a** 0 cm, **b** 10 cm, **c** 20 cm, **d** 30 cm, **e** 40 cm, **f** 60 cm, **g** 80 cm, and **h** 100 cm



(g) (h)

Fig. 2 (continued)

$$\mathrm{Scf} = \frac{L(\mathrm{cm})}{\mathrm{DD(pixel)}} \tag{1}$$

where Scf is the scaling factor which can be used to estimate the real object length in image plane, L is the actual object's length in cm, DD is an object's length in image plane measured by pixel and equal to the following expression, and $(x_1, y_{1)}, (x_2, y_2)$ are the x-y coordinates for the first and second points, respectively [13, 14]:

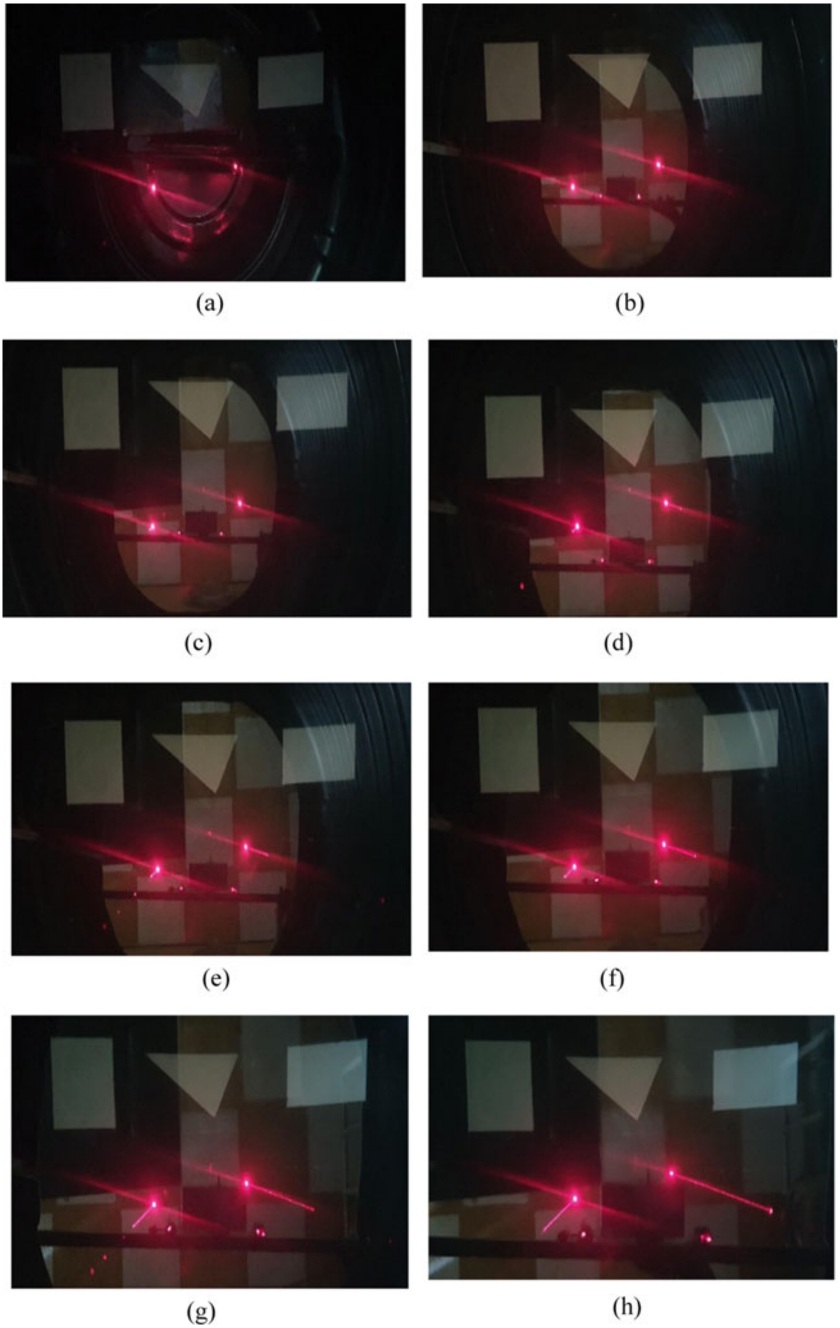
$$\mathrm{DD} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{2}$$

3 Results and Discussion

Figures 2 and 3 show the captured images for the scene which contains three different objects inside a plastic tank filled by water with and without alum, respectively.

In the visual appearance to the captured images for the two cases, a disappearance case for an object's edges began to be appeared in the turbulent water at depth equal to 80 cm and more, while a distinguished edges resulted in the distilled water images. Tables 1 and 2 show the results for algorithm's application for the used three targets.

Figure 4 demonstrates the variation of the statistical properties with water depth, a complete matching between the actual and the estimated values occurred for the case of laser two spots distance-water depth variation for the two used types of water as in Fig. 4a. Curves of all objects' properties for both water types seem to be overlapped for the two used types of water except for the triangle shape which witnessed a divergence state began to be appeared at water depth equal to 10 cm as seen in Fig. 4c, e.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 3 Captured images by using distilled water for water depth equal to: **a** 0 cm, **b** 10 cm, **c** 20 cm, **d** 30 cm, **e** 40 cm, **f** 50 cm, **g** 60 cm, **h** 70 cm, **i** 80 cm, **j** 90 cm, and **k** 100 cm

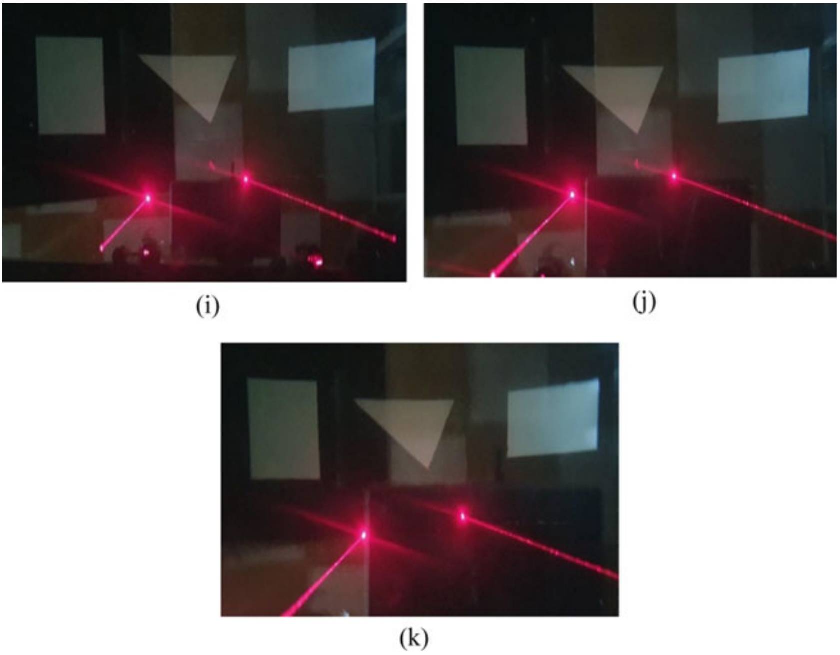


Fig. 3 (continued)

On the other hand, especially at higher levels of water depth, a distinct convergence appeared clearly between the actual and estimated values and this proves algorithm's accuracy in detecting the characteristics of underwater objects at depths exceed one meter. This can be distinguished in Fig. 5 for measuring the area of the underwater objects.

4 Conclusion

The presented work introduced a technique for measuring object's properties of three different objects immersed inside a designed system prepared for such purpose. A complete match between the actual and estimated values resulted for the laser spots distance-water depth variation for the two used sources of water. In addition to that, a distinct convergence appeared clearly upon the actual and estimated values in measuring the area of underwater object's especially at water depth that exceeds one meter and this proves algorithm's accuracy in detecting the characteristics of underwater objects especially at higher levels of water and such technique can be used in the future in detecting the target's properties in deep water like swimming pools as an example.

Table 1 Results for estimating underwater objects' properties as a function of water depth (cm) when potassium salts are added where Ds, L.Sq, L1.Tr, L2.Tr, L3.Tr, L.Rc, W.Rc, A.Sq, A.Tr, A.Rc is the estimated distance between two laser spots, square's length, first, second, third triangle's length, rectangle's length and width (cm), the area (cm^2) for square, triangle and rectangle respectively

Depth (cm)	Ds (cm)	L.Sq (cm)	L_1.Tr (cm)	L_2.Tr (cm)	L_3.Tr (cm)	L.Rc (cm)	W.Rc (cm)	Area (cm^2)		
								A.Sq	A.Tr	A.Rc
0	21	13	19.07	9.91	17.11	9.18	16.83	169	84.505	154.5
10	21	12.78	19.93	10.26	17.5	8.59	15.94	163.33	89.7037	136.92
20	21	13.05	17.95	10.69	20.69	8.56	16.62	170.3	95.9198	142.27
30	21	12.85	17.58	10.39	20.32	8.78	16.37	165.12	91.3223	143.73
40	21	13.28	17.59	10.76	20.29	8.73	17.07	176.36	94.574	149.02
50	21	13.5	17.81	10.88	20.6	8.94	17.42	182.25	96.8459	155.73
60	21	13.73	18.29	10.99	20.94	9.13	17.75	188.51	100.4155	162.06
70	21	13.92	18.68	11.8	21.41	9.57	18.32	193.77	109.96	175.32
80	21	13.91	18.42	10.88	20.98	9.4	18.19	193.49	100.1091	170.99
90	21	13.99	18.39	11.37	21.2	9.73	18.97	195.72	104.4499	184.58
100	21	14.3	18.81	11.23	21.59	9.42	19.31	204.49	105.5618	181.9

Table 2 Results for estimating underwater objects' properties as a function of distilled water depth(cm)where Ds, L.Sq, L1.Tr, L2.Tr, L3.Tr, L.Rc, W.Rc, A.Sq, A.Tr, A.Rc is the estimated distance between two laser spots, square's length, first, second, third triangle's length, rectangle's length and width (cm), the area (cm^2) for square, triangle and rectangle respectively

Depth (cm)	Ds (cm)	L.Sq (cm)	L_1.Tr (cm)	L_2.Tr (cm)	L_3.Tr (cm)	L.Rc (cm)	W.Rc (cm)	Area (cm^2)		
								A.Sq	A.Tr	A.Rc
0	0	4	3.7249	4.3681	47.4721	0.6724	10.0489	3136	1720.175625	2070.25
10	0	4.9284	1.1449	3.0276	42.25	1.9881	16.4836	3803.1889	1315.969942	3979.086
20	0	3.8025	9.3025	1.7161	10.9561	2.0736	11.4244	2992.09	903.615624	3332.753
30	0	4.6225	11.6964	2.5921	13.5424	1.4884	13.1769	3585.6144	1201.156169	3166.313
40	0	2.9584	11.6281	1.5376	13.7641	1.6129	8.5849	2365.8496	986.336836	2598.96
50	0	2.25	10.1761	1.2544	11.56	1.1236	6.6564	1827.5625	848.7957828	1959.833
60	0	1.6129	7.3441	1.0201	9.3636	0.7569	5.0625	1331.5201	653.5436603	1439.444
70	0	1.1664	5.3824	0.04	6.7081	0.1849	2.8224	975.3129	256.6404	609.1024
80	0	1.1881	6.6564	1.2544	9.1204	0.36	3.2761	992.8801	669.3034668	841.5801
90	0	1.0201	6.8121	0.3969	7.84	0.0729	1.0609	857.3184	463.545206	237.7764
100	0	0.49	4.7961	0.5929	5.8081	0.3364	0.4761	420.6601	416.9028912	327.61

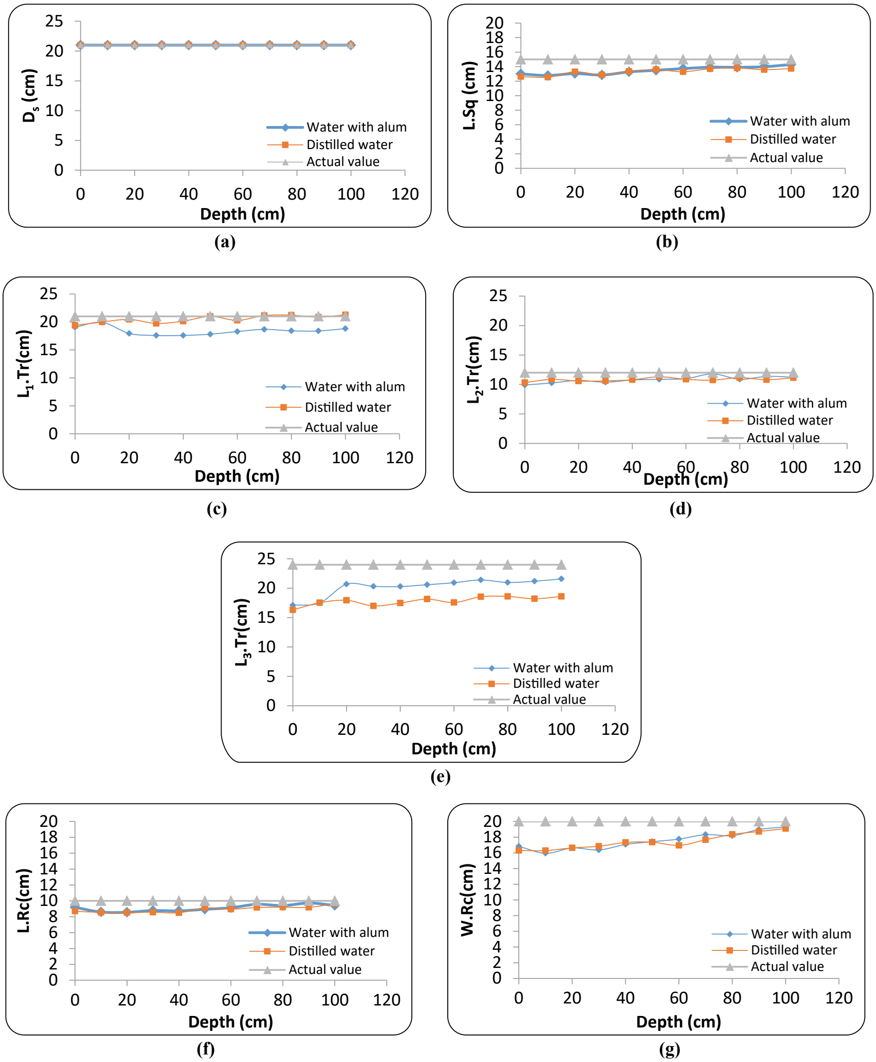
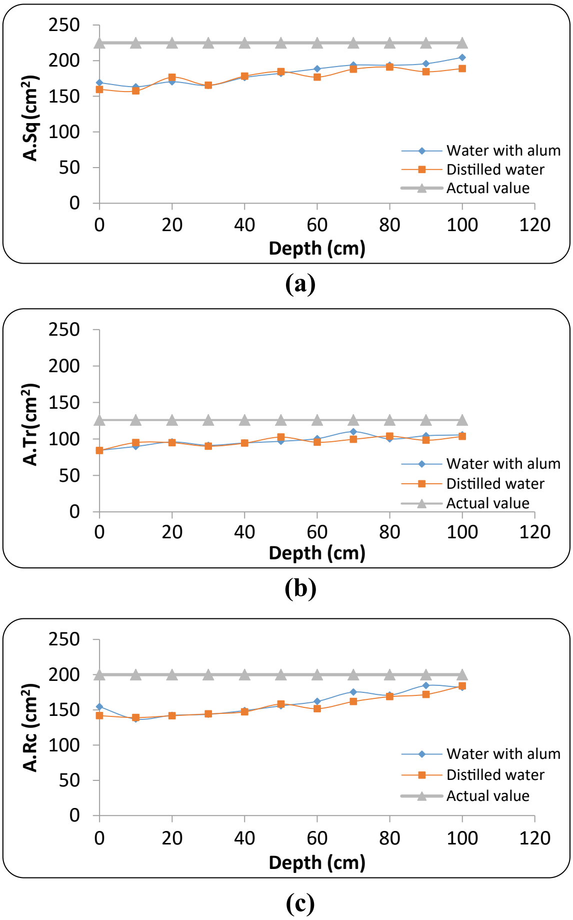


Fig. 4 Real and estimated variations of the statistical estimators with water depth (cm)

Fig. 5 Real and estimated variations of the object's area with water depth (cm) for **a** square object, **b** triangle object, and **c** rectangle object



Compliance with Ethical Standards There are no conflicts of interest besides the research did not receive any specific funding but was performed as part of the employment of the authors.

Acknowledgements The authors would like to thank Mustansiriyah University (www.uomustansiriyah.edu.iq) Baghdad-Iraq for its support in the presented work.

References

1. Murino V, Trucco A (2000) Three-dimensional image generation and processing in underwater acoustic vision. Proc IEEE 88:12
2. Yamashita A, Fujii M, Kaneko T (2007) Color registration of underwater images for underwater sensing with consideration of light attenuation. In: Proceedings 2007 IEEE international conference on robotics and automation, 10–14 Apr 2007, Roma, Italy

3. Schettini R, Conchs S (2010) Underwater image processing: state of the art of restoration and image enhancement methods. EURASIP J Adv Signal Process 3
4. Fonseca JASD, Baptista A, Martins M, Torres JPN (2017) Distance measurement systems using lasers and their applications. Appl Phys Res 9(4)
5. Mrovlje J, Vrančić D (2008) Distance measuring based on stereoscopic pictures. In: 9th international Ph.D. workshop on systems and control: young generation viewpoint 1–3 Oct 2008, Izola, Slovenia
6. Al-Obaidi FEM (2015) Quality determination for drinking water by using image quality assessment. Arab Gulf J Sci Res (AGJSR) 33(2/3)
7. Al-Shimiry I F H, Al-Zuky A A D, Al-Obaidi F E M (2021) Geometrical models for simulating the properties of underwater bodies using digital image processing. AIP Conference Proceeding, Vol. 2404, p. 030003, 2nd International Conference on Engineering and Science (ICES2021), 26-27 May, 2021, Technical Institute, Al-Samawah, Iraq. https://doi.org/10.1063/5.0068897
8. Al-asadi TA, Obaid AJ (2016) Object detection and recognition by using enhanced speeded up robust feature. Int J Comput Sci Netw Secur (IJCSNS) 16(4):66–71
9. Mohamed H, Negm A, Salah M, Nadaoka K, Zahran M (2017) Assessment of proposed approaches for bathymetry calculations using multispectral satellite images in shallow coastal/lake areas: a comparison of five models. Arab J Geosci 10:42
10. Al_Mawlawi HMJ (2011) Objects distance estimation using image processing technique based on laser pointers. M.Sc. Thesis, Department of Physics, College of Science, Mustansiriyah University, Baghdad, Iraq
11. Chybicki A (2017) Mapping south Baltic near-shore bathymetry using sentinel-2 observations. Polish Maritime Res 3(24):15–25
12. Dietrich JT (2017) Bathymetric structure-from Motion: extracting shallow stream bathymetry from multi-view stereo photogrammetry. Earth Surf Proc Land 42:355–364
13. Abbas HK, Mohamad HJ, Al-Saleh AH, Al-Zuky AA (2019) Modelling vision angles of optical camera zoom using image processing algorithm. IOP Conf Ser Mater Sci Eng 571(02117)
14. Ali SH, Al-Zuky AAD, Al-Saleh AH, Mohamad HJ (2019) Measure liquid viscosity by tracking falling ball automatically depending on image processing algorithm. IOP Conf Ser J Phys Conf Ser 1294(022002)

Digital Image Watermarking Techniques Using Machine Learning—A Comprehensive Survey

Satya Narayan Das and Mrutyunjaya Panda

Abstract Digital image watermarking is the most interesting and active field for research as it prevents unwanted access to multimedia data. The trade-off between imperceptibility, robustness, capacity and safety must be maintained for the conception of an efficient and strong digital picture watermarking system. Different studies have been conducted in order to ensure that these needs are hybridized by many domains, including spatial and transformational fields. An analytical analysis is performed on existing digital picture watermarking systems in this research. The digital information that has resulted in the request for a safe ownership of the information may recently be readily changed, reproduced, distributed and stored. The watermark solution for the authentication of content and copyright protection is quite good. This paper discusses basic concepts and features of digital watermarking, important attacks on watermarking systems, general embedding and extraction processes for watermarking marks, and important techniques for the transformation using machine learning are analysed. The objective of this paper is to provide an ephemeral study and background on the definition, and idea and major contributions of watermarking the techniques are classified according to different categories: host signal, sensitivity, robustness, kind of watermark, essential data for extraction, processing domain and applications.

Keywords Digital image watermarking · Data access · Security · Machine learning models · Watermark extraction

S. N. Das (✉)
Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India
e-mail: sndas@ieee.org

M. Panda
Department of Computer Science and Applications, Utkal University, Bhubaneswar, Odisha, India

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_39

1 Introduction

Watermarking is an information hiding technique used in digital media such as pictures, digital audio and digital video to hide private information. Copyright violations have been caused by the ease with which digital material can be exchanged on the Internet [1]. It is possible to quickly exchange copyrighted data over peer-to-peer networks, thus the content providers who produce this digital content have been particularly concerned with this. These digital contents might be labelled with watermarks to safeguard the interest of content suppliers [2]. In this study, the latest approaches used for watermarking photographs are discussed. Digital videos can also be used with image watermarking techniques [3].

The technology of multimedia is enhancing every day. Therefore, it is simple, low cost and immediate delivery without quality loss, to alter, duplicate, copy or distribute your digital imaging during communications across local networks and the Internet. The multimedia revolution has a major focus on image security and privacy [4]. In recent years, digital watermarking is a major technological improvement for the identification of proprietary information and for multimedia security for copyright holders. This method incorporates the watermark data into a media product such as image, audio, video and text and afterwards extracts it from or recognizes it as a product from a watermarked product [5]. Through insertion of watermarks data which cannot be erased or replaced by a Snap-on, the host data are safeguarded. This system ensures authentication of content, integrity checking and manipulating resistance and creates highly protected images [6].

Invisible watermarking ensures the cover image quality and beauty. Different techniques based either on a spatial or transforming domain or both have been proposed in recent years. The spatial domain approach integrates the watermark by modification of the host image pixel values. The transform coefficients are, however, adjusted in the transform domain. The changes include discreet cosine transformation (DCT), or discrete Fourier transformation (DFT) or discrete wavelets transformation (DWT) [7]. The hybrid technique combines any two or three changes to preserve a balance between imperceptibility, strength, capability and security [8].

Digital media like images, music and video are a significant means of communication in the world that are utilized increasingly in order to provide multimedia materials, making the use of different networking models easier to alter, save, distribute or reproduce the data [9]. This indicates no quality difference between the clone and the original image. Unrestricted copying and malicious manipulation, however, entail enormous financial losses and IP rights difficulties. Information hiding has thus developed into an important area of research, and watermarking is being utilized for the security of digital photographs as a data hiding method [10].

A proposal has been made for digital watermarking systems to protect content, authentication and rights of copy and distribution against unlawful copying. In this technology, the unlawful insertion of durable, unnoticeable brand into a host image is protected [11]. The algorithms for watermarking must be naked-eyed, resistant and blind to attack, which implies, for the detections and extractions of the brand, that the original image is not required. Robustness, imperceptibility, capacity, security and computational costs are important aspects of digital watermarks. Robustness, imperceptibility and ability, however, are mutually exclusive. So, these features are strongly balanced. Two domain methods are divided by the picture marking techniques:

(1) Spatial domain approaches (least significant bit replacement, spectrum spread, etc.) are simpler, but less resistant in the face of frequent signal processing attacks [12].
(2) Domain transformation methods (DWT, DCT, SVD, etc.), while computational complexity is higher than spatial domain techniques, which are more robust against common signal processing attacks [13].

Machine learning is a methodology for determining predictions from previous data. It is a method that enhances the watermark detection rate after being assaulted and contains multiple ways and patterns of problem recognition for distinct classifications [14]. It also allows more automation through the replacement of long-term activity with automatic procedures utilized for the improvement of accuracy and efficiency in the process of knowledge engineering. As stated below, in general, four types of assaults on digital watermarking systems.

(1) **Active Attacks**

Hacker attempts to remove watermarks in active assaults. They are designed to distort watermarks before they are recognized. This shows the copyright, copy control, etc., issues.

(2) **Passive Attacks**

Hackers have only tried to identify and recognize watermarks in passive attacks. In passive attacks, no harm or deletion is done during this attack.

(3) **Falsification Attacks**

Hackers implant a new and valid watermark in falsification attacks rather than removal. In this attack, the data may simply be manipulated and the image can be distorted.

(4) **Collusion Attacks**

Hackers have the same intension in collusion attacks as for the active attacks, however, have utilized somewhat different approaches. The hacker utilizes the same data instances to build the new copy without a watermark.

The quality of digital audio, pictures and video signals is higher than that of their analogue equivalents and offers little diverse advantages over analogue media. Editing is easy because it is possible to modify just the places to be modified. The reproduction data is straightforward without quality loss [15]. A digital media copy is the same as the original. The integration of the copyright label or digital signature, termed watermark, which completely characterizes the person who applies and so identifies it as intellectual property, that is one means of protecting multimedia data against illegally logging and transmissions [16]. In 1993, Tirkel introduced two watermarking techniques for embedding the watermark in pictures and coined the phrase ‘digital marking’.

Digital watermarking is a technology that offers and guarantees digital media security, data authenticity and copyright protection [17]. Digital watermarking is the process of incorporation into the digital media, e.g. text, image, audio and video of data termed a watermark that allows a later detection or extraction of watermark to claim the object [18]. A digital watermarking technique incorporates information into a document or media directly. For example, copyright information, ownership information, time stamps and the legitimate recipient may be integrated. Digital watermarking cannot, by itself, prevent documents from being copied, amended and redistributed [19]. However, failure to encrypt and copy protection allows watermarking to trace a document back to its legitimate owner and to an illegitimate usage point [20].

Images will be changed into frequency domain in the converted domain, and then data will be incorporated in converted coefficients. It is also available in different sectors, such as ownership assertion, copyright protection, broadcast surveillance, ID card security, fingerprinting, content labelling, copy control and manipulator detection, authentication of the contents, usage control, integrity verification, medical security and content protection according to the different watermarking characteristics [21]. In a single watermarking system, sometimes numerous applications might be coupled. However, all applications cannot be placed into one scheme since different applications require distinct watermarking system characteristics to a different degree [22]. Different qualities or requirements of watermarking also occur and lead to diverse design difficulties depending on the type of watermarking applications and purposes. Recent watermarking technology has been implemented and achieved with the combination of spatial domain and frequency area in terms of robustness, capabilities for perceptibility, security and low complexity [23].

Postage stamps and banknotes have been usually watermarked historically. In fact, when printing banknotes, the currency watermark is still used today. A digital watermark can be seen or unseen. The luminous logos that often appear in the corner of films and pictures to avoid copyright infringement are an example of a digital visible watermark [24]. However, by simply cutting the media or writing the logos, these visible water markers can be targeted and eliminated. The field of digital watermarking is then concentrated mostly on including invisible watermarks that function exceptionally changing the media contents. As it cannot be seen, a robustness property has to be established that ensures a survival of the watermark data when the image is changed. The watermarking process is the integration into media like images, audio or video of particular data. This embedded information, known as the watermark, may later be removed and utilized to support ownership from the multimedia material [25].

1.1 Issues and Challenges in Watermarking

In watermarking research, there are many technical hurdles. Research is highly fascinating due to the trade in robustness and imperceptibility. To achieve imperceptibility, the watermark should be added to the high-frequency components of the original signal [26]. On the other hand, only the low-frequency components can be added to the watermark for stability. The watermarking approach can therefore work when the low-frequency components of the original signal are employed as host for insertion of watermarks [27].

In the era of technology, digital multimedia editing, manipulation and copying are now commonplace and becomes a severe problem. One of the answers to this problem is digital watermarking [28]. The approach is used to include patterns of bits into the digital data in such a way that the actual data is not hindered. It, therefore, retains its trustworthiness. A wide number of applications such as copyright protection, fingerprinting, copy safety and medical applications can be employed for digital watermarking [29]. We have identified several challenges and problems that arose throughout the application of this methodology.

The first thing to do is keep the balance between imperceptibility, robustness and capacity because one factor has a detrimental influence for another. A hot watermark should be incorporated into high-frequency components to obtain good imperceptibility, whereas robustness in a low-frequency component should be achieved. The human visual system is another problem. Only blue colour is less sensitive to hidden watermarks in the RGB colour pictures. Another is delicate watermarking, and the recovery of contents from image is a difficult matter [30]. As with a delicate watermark, the watermark is slightly distorted. One is the size of the payload; how much information it contains and is the size of the payload. As the payload is larger, the imperceptibility is compromised. Therefore, it is a matter of maintaining balance. The next problem is spatial domain robustness. Like the space domain, the pixel values are changed. Different attacks like compression JPEG, high pass filters, low-pass filtration, cuts, etc., are scarcely resistant. Another issue is the cost of calculating watermarks that should be minimized [31]. The next issue is the false positive rate of the digital marking system which is vital to be considered. Another difficulty is the construction of a universal method that is robust to multiple kinds of attacks for all digital media.

The above difficulties can be handled by incorporating certain ownership details inside the multimedia data, which can later be retrieved to establish their ownership. This principle is implemented in bank notes that are implanted with a watermark to verify the authenticity of the currency. In multimedia, digital contents, the same 'watermarking' notion can be utilized for authenticity assessment [32]. In watermark, the original data contains a secret and imperceptible signal, which remains present, so long as the observable content quality is satisfactory. In case of several ownership claims, by removing the watermark from the watermarked contents the owner of the original data can claim the ownership [33].

2 Literature Survey

A method based on the SVD and DCT Walsh hybrid transform is proposed for colour-picture watermarking by Natu et al. [1]. This approach uses low-frequency coefficients to insert the watermark instead of the DCT coefficients of the mid-frequency transformation. The experimental results demonstrate significant robustness against noise, histogram, compression, and crop attack. This approach offers great strength. Another solution proposed by Assini et al. [2] combines DWT, DCT and SVD to secure medical images for the improvement of system robustness. In the medical image here, an invisible watermark image is inserted. The DWT 3-Level is applied to the image of the medical host. The DWT coefficients of the medical host picture use the high-frequency sub-bands in this context. The host picture is then subject to DCT and SVD alterations. The technology guarantees greater noise imperceptibility and robustness (gauzy and salt-and-pepper sounds).

In a model proposed by Takore et al. [3], LWT, DCT and SVD are coupled to use the border detection to choose the optimal area to integrate the watermark. In order to maintain the balance between robustness and imperceptibility, the particle swarm optimization (PSO) technique is applied. The testing results reveal that the system is robust against several attacks and also retains the visual quality of the host. A hybrid DCT-SVD-based frequency domain approach for increased capacity is proposed by Jain et al. [4]. The Arnold transforms this tough watermarking approach to texture the aqua mark logo. The DCT is applied on the host image first of all two-dimensional (2D), and the SVD of the watermark logo is taken. Different weights are set to minimize the distortion of the host image through lowered singular values. Two domain processes such as discrete Fourier transform (DFT) and DCT are combined to secure copyright information that is proposed by Hamidi et al. [5].

The adaptive watermarking picture approach based upon a neural network was presented by Oueslati et al. [8] to secure medical data. The watermark is introduced within the coefficients of the DCT-based transform domain covered by blocked photography. This will strengthen the watermark and make it less likely to attack certain sorts. A survey on digital watermarking image techniques was done by Potdar et al. [9]. This methodology is based on many areas in which data is incorporated. A visible and invisible watermark is incorporated in the multimedia object in watermarking procedures. A security key guides the embedding procedure.

In order to generate virtually obvious deformation of the watermarked image, Tripathi et al. [10] developed digital watermarks leverage the features of discrete cosine transformation (DCT) and discrete transformations. These strategies employed a unique mechanism to disseminate, embed and extract the watermark. Wajid et al. [11] proposed resilient and invisible watermarking by employing the whole neural network counter-propagation. A trade-off between robust and imperceptible watermarking features is always provided by various strategies. The neural network full counter-propagation is used to train numerous greys or colour pictures to make watermark pictures. Yusof et al. [12] proposed a digital watermark model using wavelet transform for digital photos. The wavelet transforms an image into a series of confined elements which may be reconstructed without error to recreate the original image. Watermark is incorporated in the image by means of the quantizing process in the band pass wavelet coefficients of high amplitude.

Alizadeh et al. [13] suggested a model for image watermarking on the basis of Q-learning matrix factorization that is a strengthening learning model. Here, Q is used to identify the proper embedding host blocks using the test and error method. It gives a better result than random embedding for imperceptibility and strength. The performance against several assaults is not indicated in this model. Another model is suggested by Mehta et al. [14] using the genetic algorithm supporting vector machine. Important regions are picked, here with fumigating entropy and prominent low-frequency regions are created using the support vector regression model from these important regions, which boosts resilience significantly over prior methods. Here, a genetic algorithm is used to calculate the watermark scaling factor.

The most popular usage of optimization algorithms proposed by Dubey et al. [15] is for discovering right embedding functions or the proper integration block, such as genetic algorithm, swarm optimization, ant colony optimizations and firefly algorithms. The neural network is now more popular than other methods for learning the machine. Guo et al. [16] proposed a system that adds a watermark onto an image using the predictive quantity of prepared data streams for direct level selection. This data stream is selected to look like quantization noise for the resultant image. There is a difference in this approach in which a watermark in the form of a dithering matrix is utilized to dither an image. There are, however, limitations to even these arrangements. The fundamental issue is that the signals, particularly quantizing and geometrical attacks are sensitive to transmission such as crop. They also degrade an image in a way which can be predicted and dithered. During the original phase, the watermarking technology was designed by the application of a special domain which had become straightforward to implement and relatively less complex.

A DWT spiking neural network (SNN) with low-time complexity is developed, and the extraction problem is addressed as an SNN optimization issue analysed by Kazemi et al. [18]. A lot of work is done based on a network of neurons which, in addition to known frequency fields like DCT or DWT that have been merged with contourlet transformation, Kurtose coefficients and YCbCr spaces analysed by Yadav et al. [20]. Vatsa et al. [21] introduced biometric watermarking image algorithms that incorporate the face image into the fingerprint. The watermarking process is based on a discrete wavelet transformation (DWT) and support vector machine (SVM). The experimental results have been proved to be robust and to be able to withstand geometrical and frequency assault. SVM integration enhanced facial acknowledgement by 10 per cent. Wu et al. [22] proposed a spatial domain-based digital watermarking method based on SVM. In SVM training, watermarking just uses 128 bits. This approach will modify blue pixel channels at the same time in order to incorporate the watermark bits. Watermarks are integrated into a spatial domain without the need for an original image and are extracted straight from a watermark image. The research findings reveal the strong PSNR of watermarked images and a low extraction error rate in the suggested approach.

Li et al. [23] proposed the invariant watermarking approach rotation, scaling and translation (RST) using SVM and picture moments. SVM is used to learn the geometric image pattern in six combined low order photo moments in watermarking methods for estimating RST transform parameters. The experimental results suggest that JPEG compression, noise and geometric attack can be resisted by the system. The model uses a fuzzy support vector machine watermarking system that uses spatial domain (FSVM). The 8 * 8 block of the cover picture is sub-blocked as input vectors through SVM throughout the embedding process. The sub-block of the image has been separated into a strong and weak texture. In the cover image, the strong texture information is integrated. The approach has been demonstrated that the solidity of the FSVM method against major attacks is superior than that of an SVM approach.

The watermarked algorithm using SVM based on the colour picture was proposed by Jain et al. [24]. The watermark is included into the original image's discretionary wavelet domain and extracted by a training support vector machine during the embedding phase. Furthermore, the method uses an impulse coefficient to lower the error and boost the learning rate. The test results have revealed that the approach against signal processing attacks is not perceptible. For most attacks, however, the PSNR value is below 27 dB.

Ramamurthy et al. [25] suggested a robust, neural network- and fuzzy-like watermarking picture system. The methodology compares the use of DWT domain quantization in the back propagation neural network (BPNN) and dynamic fuzzy inference system to embed watermark into the host image (DFIS). The experimental outcome shows a durable and unnoticeable watermarking strategy for assaults. The robust and blind image watermarking algorithm for the protection of copyrights for photographs was proposed by Jagadeesh et al. [26]. DWT uses the integration watermarking approach on the basis of the SVM. The test results suggest that the technique for diverse attacks is secured and sturdy. The NC and PSNR values for the majority of the attacks, however, are less than 0.9711 and 35 dB.

A strong blind watermarking method was proposed by Vafaei et al. [27]. The watermarking approach employs the discrete wavelet transformation domain of the neural networks. The neural networking technology is utilized to maximize the image's strength. The experimental results suggest that the approach is robust and unaware of different attacks. A new approach of picture watermarking in discreet wavelet transform domain utilizing a support vector machine was proposed by Jagadeesh et al. [28]. Even after different image processing attacks, the integration method is employed to extract the watermark from the watermarked image. The testing results suggest that the algorithm provided is safe and sturdy for various picture attacks.

Yahya et al. [29] created an SVM stage categorization information security model. The embedding technology uses image steganography which hides data beneath a cover image in the DCT field. The methodology developed a novel model using the HVS and the LSB shifting techniques known as Stego SVM-DCT-LSB, to maintain imperceptibility and to enhance the resilience of stego-images. The technology of image watermarking using neural networks was suggested by Zhang et al. [30]. The watermarking technique is used as a type of communication to hide information in the form of pictures. The experimental results demonstrate that the associative memory attraction basin determined watermarking ability.

Shi et al. [31] presented a new circulation-based, non-overlapping SVD-backed colour embedding methodology, for the concealment of essential image information. The cover picture is decomposed in small watermarks by the watermarking procedure, and then watermarks are integrated in one block with circulation. The results of the experiment reveal that the system is robust against various image attacks. Thai et al. [32] presented picture classification methodology that would use the SVM and artificial neural network. The approach divides the picture into sub-images and categorizes each sub-image into ANN, then the SVM compiles the entire ANN's classified output. The experimental results demonstrate the technique's practicality. It is obvious from the foregoing that watermarking images in a transforming domain have been found to be very resilient, imperceptible, capable and secure. The digital watermarks may be beneficial for several applications such as affirmation of ownership, finger printing, copying or monitoring, telemedicine, e-governance, e-commerce, forensic media, artificial intelligence and health care.

Kim et al. [33] examined briefly the topic of digital watermarking. In the less relevant bits of pixels in the area of the image contours, the authors offered a description of a method for inserting a watermark. The watermark is easily removed as it relies on changes of the least significant bits. Moreover, its approach is limited to images, in that the watermark is inserted into picture areas on the contour edge. A binary bit's mask with the LSB of each pixel is compared against the addition or subtraction. If the mask bit is equal to LSB, the random amount will be added, otherwise, it will be subtracted. In order to determine, whether this conforms to the original sequence of additions and subtractions, the watermark is deducted by first calculating the difference between the original and watermarks, then checking the difference sign, pixel by pixel. However, the idea is that high-frequency noise is used to ensure a certain robustness in low-pass filter, which does not have perceptual relevance. This technique does not take into account the issue of collusion.

3 Proposed Model

Technologies of spatial domains do not offer a high level of attack robustness or manipulation. In this regard, multimedia safety has been paid attention to transform domain techniques [29]. However, in recent years', hybrid domain approaches are favoured because to the restricted payload capacity of transform domain approaches. Two or more picture alterations for watermarking are employed in hybrid domain processes. The extensive forms of transform domain methods are these hybrid domain approaches. These methods give the multimedia data a greater imperceptibility and robustness and are generally utilized for the protection and protection of multimedia.

Hybrid digital watermarking is a host picture method which uses two or more DCT, DFT, DWT and SVD domain approaches. Any lightweight encryption approach encodes the watermark image, and then divides the encrypted watermark image into n-blocks. The encrypted block watermark picture will thereafter be placed in the host picture. The host image n-places here are chosen randomly for the inclusion of the encrypted watermark in n-blocks. Then, the reverse change is made and eventually, the watermarked image is achieved. The extraction of the watermark is done in the opposite way. The watermark embedding framework approaches is shown in Fig. 1.

The digital picture watermark technology integrates data/image into an image so that the data cannot be deleted or destroyed easily. There are various design requirements for efficient watermarking. In addition, the combination of different domain transform algorithms is possible in hybrid watermarking. During transmission over the Internet or the personal network, it is easy to reproduce, transfer and distribute a digital image. Therefore, digital image watermarking adds host media information in order to prevent unauthorized access. The system should, therefore, be effective. For efficient hybrid digital image watermarking processes, some parameters must be addressed. Among other needs, imperceptibility, strength, ability and safety are the four fundamental needs (Fig. 2).

Based on the problems identified, there is a necessity to overcome the issues by providing the solutions to the observed drawbacks. The proposed model requires to

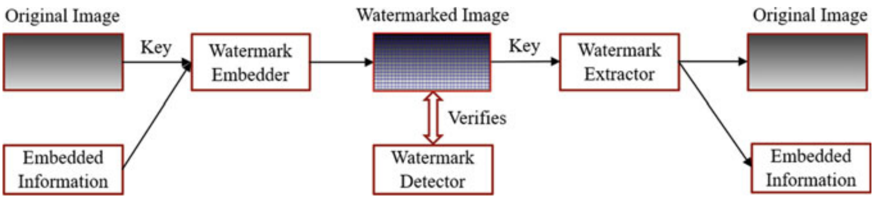


Fig. 1 Digital watermarking process

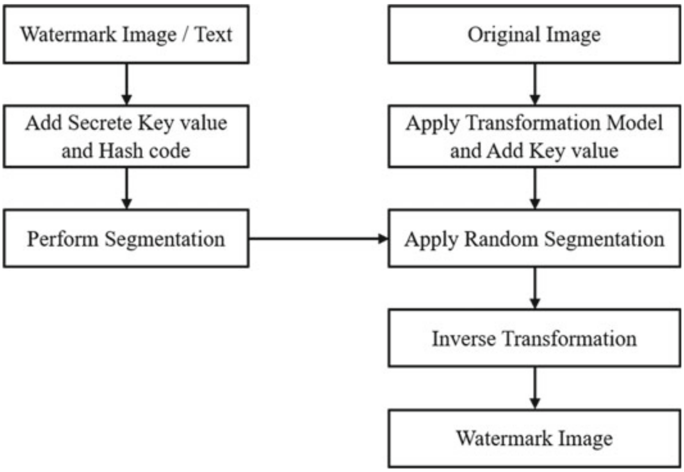


Fig. 2 Watermark embedding process

understand image watermarking requirements depending on their applications. The initial stage in the development of algorithms for various watermarking applications is the clear knowledge of these requirements. Then, how to improve the resilience of picture watermarks need to be considered and implement watermark algorithm to see how they fulfil various application requirements and generic watermarking requirements. Machine learning techniques can be used to improve the accuracy levels of the model for easy and accurate image watermarking at sender side and content extraction at the destination side.

4 Conclusion

Images provide an essential part of multimedia data. Authentication of images is a difficult issue because of traffic on the Internet. It is important to provide imperceptibility, robustness and improved data inclusion capacity alongside image data security. A major area to ensure these problems is the digital watermarking of images. Digital watermarking may be utilized in any sector where multimedia data must be protected with guaranteed safety and authenticity for identification, annotation and copyright purposes. Transform domain technology watermarking takes advantage of the spatial and frequency information in various resolutions to increase resilience. Inclusion and extraction process of the watermark, gain parameters, noise changes and watermark size are the key to the performance of the existing watermarking techniques. This paper provides a brief survey on current watermarking procedures that ought to be upgraded in order to resolve several problems. Therefore, future researchers must combine artificial neural network and machine learning methods in the cross-transform domain with enhanced resilience and high image data security, and improving imperceptibility and integration capabilities. We intend to build and deploy a novel technology of digital watermarking for images utilizing machine learning technology based on space and frequency. The many watermarking approaches are used to learn about the performance and the state-of-the-art parameters. The watermarking area can be further explored using the machine training methods to improve the accuracy rate.

References

1. Natu S, Natu P, Sarode T (2017) Improved robust digital image watermarking with SVD and hybrid transform. In: Proceedings of the international conference on intelligent communication and computational techniques (ICCT), Jaipur, India, December 2017.View at Publisher Site, Google Scholar, pp 177–181
2. Assini I, Badri A, Badri A, Safi K, Sahel A, Baghdad A (2018) A robust hybrid watermarking technique for securing medical image. Int J Intell Eng Syst 11(3):169–176
3. Takore TT, Kumar PR, Lavanya Devi G (2018) A new robust and imperceptible image watermarking scheme based on hybrid transform and PSO. Int J Intell Syst Appl 10(11):50–63

4. Jain P, Ghanekar U (2018) Robust watermarking technique for textured images. In: Proceedings of the 6th international conference on smart computing and communications (ICSCC), Kurukshetra, India, December 2018. View at Publisher Site, Google Scholar, pp 179–186
5. Hamidi M, Haziti ME, Cherifi H, Hassouni ME (2018) Hybrid blind robust image watermarking technique based on DFT-DCT and Arnold transform. Multimedia Tools Appl 77(20):27181–27214
6. Craver S, Memon N (1998) Resolving rightful ownership with invisible watermarking techniques: limitations attacks and implications. IEEE Trans 16(4):573–586
7. Maurer A, Hersch M, Billard AG (2005) Extended hopfield network for sequence learning: application to gesture recognition. In: Proceedings of 15th international conference on artificial neural network, pp 493–498
8. Oueslati S, Cherif A, Solaimane B (2011) Adaptive image watermarking scheme based on neural network. IJEST 3(1)
9. Potdar VM, Han S, Chang E (2005) A survey of digital watermarking techniques. In: IEEE International conference on industrial informatics
10. Tripathi S, Jain RC, Gayatri V (2006) Novel DCT and DWT based watermarking techniques for digital images. IEEE
11. Wajid SK, Jaffar MA, Wajid R, Mirza AM (2009) Robust and imperceptible image watermarking using full counter propagation neural networks. In: International conference on machine learning and computing, vol 3
12. Yusof Y, Khalifa AO (2007) Digital watermarking for digital images using wavelet transform. IEEE International
13. Alizadeh M, Sajedi H, Babaali B (2020) Image watermarking by Q learning and matrix factorization. In: Proceedings of the international conference on machine vision and image processing (MVIP), Qom, Iran, February 2020. View at Publisher Site, Google Scholar
14. Mehta R, Gupta K, Yadav AK (2020) An adaptive framework to image watermarking based on the twin support vector regression and genetic algorithm in lifting wavelet transform domain. Multimedia Tools Appl 79(25–26):18657–18678
15. Dubey G, Agarwal C, Kumar S, Singh HP (2020) Image watermarking scheme using cuckoo search algorithm. Advances in data and information sciences, lecture notes in networks and systems, vol 94. Springer, Singapore
16. Guo Y, Li B-Z, Goel N (2017) Optimised blind image watermarking method based on firefly algorithm in DWT-QR transform domain. IET Image Process 11(6):406–415
17. Chen Y-H, Huang HC, Lai CH, Chang TY (2020) An image watermarking approach based on artificial fish swarm algorithm. In: Proceedings of analysis of watermarking framework for color image through, Kyoto, Japan, March 2020, pp 46–50
18. Kazemi MF, Pourmina MA, Mazinan AH (2020) Analysis of watermarking framework for color image through a neural network-based approach. Complex Intell Syst 6(1):213–220
19. Kazemi MF, Pourmina MA, Mazinan AH (2020) Novel neural network based CT-NSCT watermarking framework based upon Kurtosis coefficients. Sens Imaging 21(1)
20. Yadav AK, Mehta R (2019) Local coupled extreme learning machine based image watermarking using DCT in YCbCr space. In: Proceedings of the amity international conference on artificial intelligence (AICAI) Dubai, UAE, Feb 2019, pp 527–532
21. Vatsa M, Singh R, Noore A (2005) Improving biometric recognition accuracy and robustness using DWT and SVM watermarking. IEICE Electron Express 2(12):362–367
22. Wu J (2009) A RST invariant watermarking scheme utilizing support vector machine and image moments for synchronization. In: IEEE international conference on information assurance and security
23. Li L, Ding W-Y, Li J-Y (2010) A novel robustness image watermarking scheme based on fuzzy support vector machine. In: IEEE pattern recognition and intelligence system
24. Jain YK, Tiwari S (2011) An enhanced digital watermarking for color image using SVM. Int J Comput Sci Inf Technol 2(5)
25. Ramamurthy N, Varadarajan S (2012) Robust digital image watermarking scheme with neural network and fuzzy logic approach. Int J Emerg Technol Adv Eng 2(9)

26. Jagadeesh B, Rajesh Kumar P, Chenna Reddy P (2013) Robust digital image watermarking scheme in discrete wavelet transform domain using support vector machine. Int J Comput Appl 73(14)
27. Vafaei M, Mahdavi Nasab H, Pourghassem H (2013) A new blind robust watermarking method based on neural networks in wavelet transform domain. World Appl Sci J 22(11)
28. Jagadeesh B, Rajesh Kumar P, Chenna Reddy P (2014) Digital image watermark extraction in discrete wavelet transform domain using support vector machine. ACEEE Int J Recent Trends Eng Technol 11
29. Yahya S, Hussain HS, Fakariah Hani MA (2015) DCT domain stega SVM-shifted LSB model for highly imperceptible and robust cover image. In: International conference on computing and informatics, vol 43
30. Zhang F, Zhang H (2004) Image watermarking capacity using neural network. In: IEEE international conference of web intelligence
31. Shi Z, He L (2010) Application of neural network in medical Image processing. In: International symposium on networking and network security
32. Thai LH, Hai TS, Thuy NT (2012) Image classification using support vector machine and artificial neural networks. Int J Inf Technol Comput Sci 32(38)
33. Kim M, Li D, Hong S (2013) A robust and invisible digital watermarking algorithm based on multiple transform method for image contents. In: Proceedings of the world congress on engineering and computer science, vol I. WCECS, 23–25 Oct, San Francisco, USA

Finding Shortest Path in Road Networks Based on Jam-Distance Graph and Dijkstra's Algorithm

Sarah Fouad Ali, Musaab Riyadh Abdulrazzaq, and Methaq Talib Gaata

Abstract Finding the shortest path in the road networks is an urgent issue for the vehicle driver to reach their destination in the shortest time which leads to consuming less fuel. In this article, a framework has been suggested to direct vehicle drivers to their destination in a shortest time taking into account the actual distance and the traffic jam of the road network. The framework consists of three stages: data clustering stage, merging stage, and determining of shortest path stage. The experimental results illustrate that the proposed framework has similar accuracy (98%) to find a shortest path as compared with existing work; on the other hand, the running time of the RN-CMS was the shortest running time; it has been achieved (1.05 s) for 100 inquires.

Keywords Road network · Traffic jam · Micro-cluster · Shortest path · Trajectory of vehicle

1 Introduction

The growing use of global positioning system (GPS) devices and autonomous navigation systems enables for the real-time tracking and processing of large streams of track data. This enables real-time analyzes of mobility data, which in turn have the potential to significantly enhance and monitor traffic flow issues, track real-time incidents, etc. Even though some of the existing methods distinguish past motions of moving objects from historical path data, whether addressing the problem of identifying clusters of moving objects from data streams or addressing the problem of recognizing how movement behaviors unravel over time, e.g., they fail to capture typical traffic routes or traffic jams—such methods fail to recognize how movement behaviors unravel over time [1]. Trajectory data of moving object is generated from a variety of sources, including GPS-equipped cars, mobile phones, online "check-in"

S. F. Ali (✉) · M. R. Abdulrazzaq · M. T. Gaata
Department of Computer Science, Mustansiriyah University, Baghdad, Iraq
e-mail: sarah.almusawi92@gmail.com

M. T. Gaata
e-mail: dr.methaq@uomustansiriyah.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_40

and "check-out" data, geo-tagged social media posts, and radiofrequency identification. RIFD [2]. A trajectory is often described by a sequence of time-stamped geographic locations [3]. A moving object can be derived as a trajectory of user, animal, a vehicle, and natural phenomenon [4]. Clustering is one of the most significant analytical methods; it tries to group data into clusters with a high similarity between cluster members and a low similarity between members of different clusters [5]. Clustering stream data is more difficult than clustering traditional data because of the following issues: (a) single-pass processing due to continuous data arrival, (b) unbounded size of data stream and limited memory space and time, and (c) evolving data with a changing model underlying the data stream. Thus, the clustering algorithm should be capable of detecting such changes [5, 6]. In the big data era, the massive traffic data presents significant challenges for traditional hotspot identification and spatial clustering approaches. Traditional region-based techniques cannot provide the level of detail needed for route tracking. The data mining of road networks is particularly concerned in the trajectory clustering of moving objects because of its functional importance in many applications, and most approaches to this problem depend on distance measurements and have a number of drawbacks, including inexact clustering, costly processing, and the inability to handle high-dimensional trajectory data [7]. Many approaches have been done to clustering trajectory, density-based clustering is one of the most popular clustering approaches, and this method is based on assigning the region to the cluster that is nearest to it [8, 9].

2 Related Works

There are mainly two approaches to handling difficult challenges in road networks, such as traffic flow prediction and shortest path finding: The first is based on vehicle trajectories, while the second is based on GPS stream data. Silva et al. [10] suggested an incremental algorithm for finding and maintaining density-based clusters in trajectory data streams called CUTiS; it is provided by use of trajectory data streams to track moving objects and iteratively preserve sub-trajectory clusters. Makariye [11] suggested new framework called "toward online shortest path" that uses the Dijkstra algorithm to allow drivers to gather the shortest path, as well as other paths and traffic updates.

Shein et al. [12] proposed an incremental sub-trajectory (iSTCM). The incremental clustering based on micro-group (iSTCM) framework is a clustering method that employs a micro-group structure for maintenance. Mao et al. [13] presented that online clustering of streaming trajectories (OCluST) is a sliding window model-based online approach for clustering streaming trajectories. It has a micro-clustering feature that groups and summarizes the most recent sets of trajectory line segments. Shein et al. [14] propose an incremental crowd discovery method over a changing data stream. In this approach, first, propose micro-group-based clustering to identify the group, and then incrementally detect the crowd structure form. Puntheeranurak et al. [15] suggest two stages of micro-group-based clustering method over a changing data stream to decrease computational time complexity. Micro-group-based clustering in the first stage is to represent moving object groups, then incremental identification of traveling companions from these possible moving object groups in the second stage. Zhang et al. [16] propose an algorithm called "Dijkstra's-DB-SCAN" which is a novel network-based density clustering technique that combines a modified Dijkstra's shortest path algorithm with density-based spatial clustering of applications with noise (DB-SCAN), and it is designed to handle with traffic accidents. Silva et al. [17] suggest NET-CUTiS which is online incremental clustering method for analyzing data streams from moving objects whose routes are constrained by a road network. Using trajectory data streams, this approach tackles the problem of recognizing and tracking the evolution of clusters of trajectories over road networks. John et al. [18] suggestion is a clustering method is density-based clustering. According to the needs of this work, the spatial and temporal components are sorted, divided, and combined. Moreover, the spatial distance between moving objects is calculated using non-Euclidean distance. Wang et al. [19] suggest first which creates a traffic model for the urban road network. The general linear model for the optimum path planning issue is then developed using the linear programming technique. Finally, the Dijkstra algorithm is improved. The best path is the one that takes the least amount of time while taking into account actual traffic limitations and vehicle time delays at intersections. Li et al. [20] presented a two clustering technique for detecting urban hotspots using taxi trajectory data. During the first phase, spatio-temporal hierarchical density-based spatial clustering of applications with noise (DB-SCAN) is used to cluster trajectory points with spatial and temporal features, and in the second phase, the concept of region growth is added to filter noise. Li et al. [21–24] suggest a study of the definitions of urban built-up areas in numerous countries, as well as a study of their essential characteristics. The justification for using taxi trajectory data to define the urban built-up zone is explored, as well as the benefits of using it and to group the point of the trajectory.

3 The Proposed System

In this study, the road network framework (RN-CMS) has been proposed to find the shortest path in term of distance and traffic jam for vehicle drivers from their current locations to their destinations. The RN-CMS framework includes three main stages: the data clustering stage, the merging stage, and the shortest path finding stages as shown in Fig. 1.

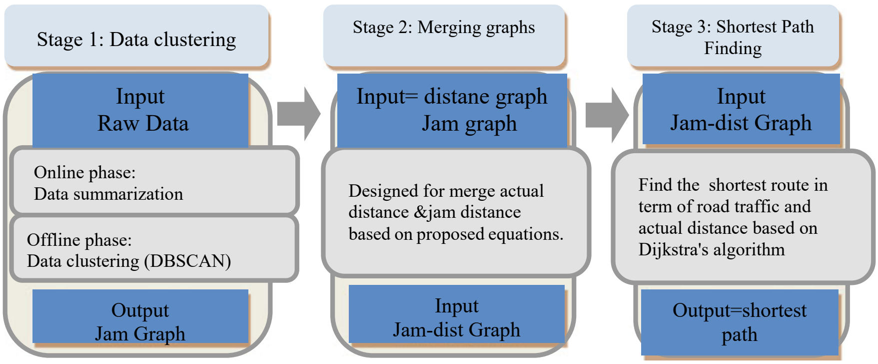


Fig. 1 Main stages of RN-CMS framework

3.1 The Data Clustering Stage

The first stage of the RN-CMS is the data clustering; it consists of two phases: the online and the offline phases. The online phase has been dedicated to summarize GPS data in the covered networks, due to the high speed and massive size of data generation. The data structure which has been supported in this phase is derived from the notion of micro-cluster in the BIRCH clustering algorithm, as defined in the following form:

$$\mathrm{MC}_j = (N, \mathrm{Ls}_{i1}, \mathrm{Ss}_{i1}, \mathrm{Ls}_{i2}, \mathrm{Ss}_{i1}, \ldots, \mathrm{Ls}_k, \mathrm{Ss}_k) \tag{1}$$

where N is the number of vehicles in MC_j and Ls_i, Ss_i are the linear sum and the square sum of vehicle location [latitude, longitude]. The main steps of the online phase are illustrated in Fig. 2.

The offline phase starts when any vehicle driver asks the system to determine the shortest route to their destination based on distance and traffic jam. The DB-SCAN clustering method has been adopted in this phase to cluster the $\mathrm{MC_c}$ instead of the raw data due to its massive size. The main outcomes of the offline phase are the jam graph which represents the traffic density in each road intersection since the $\mathrm{MC_c}$ in the connecting roads has been ignored to minimized the running time. Figure 3 explains the main steps of the DB-SCAN (offline stage); furthermore, the jam graph is shown in Fig. 4.

3.2 Merging Graphs

The goal of this step is to merge the jam graph, which is the result of the data clustering stage, with the dist graph to create a jam-distance graph in order to identify a typical

```
        Algorithm 1. Data Summarization (online phase)
Input:          Vehicle dataset={ Vh1, Vh2, …, Vhi, …. , Vhn }
OutPut:        MCc= { MCc1, MCc2, …, MCcj, …. , MCcm }
Parameter: Distanc 'd' ,
      Vhmin            /* the minimum number of item in CF */
Begin
For each pattern Vhi in Vehicle dataset
        Find the nearest MCcj to pattern Vhi
            If Dist(MCcj, Vhi) <= Dist then
               Update MCcj information according to vhi pattern;
            Else
                Create a new MCcnew for Vhi ;
      If MCc size exceeds the available memory space then
                  merge the most similar MCc
     end for
     for every MCc
   If the number of patterns in MCkn<Vhmin then
                 Discard Ckn;            /* noisy patterns*/
    end for
end.
```

Fig. 2 Main steps of the online phase

```
          Algorithm 2. The DBSCAN clustering algorithm
Input:MCc={ Mcc1, Mcc2, …, Mccj, …. , Mccm }, MinMccs, ε
Output: The Jam- Graph for the road network
Begin
All MCc set as unvisited;
   Repeat
   Choose randomly an unvisited  Mcci  micro-cluster;
   Set  Mcci  that choose as visited;
   If   εneighbor of a  Mcci  at least has  MinMccs  items
      A new cluster Clr  Create, and a Mcci added to Clr;
      Let the Nghi be as a  set of items in the εneighborhood of MCci;
         for each Pt 'point'  Mcck  in Nghi
          if  Mcck  is still unvisited
              set  Mcck as visited;
          if the εneighbor of the Mcck  at least has MinMccs
                   add those Pts to Nghi;
          if Mcck isn't belong to any cluster yet, add Mcck to Clr;
        end for
          output  Clr;
    else set  Mcck as noise;
   until all of the items have been visited
end.
```

Fig. 3 Steps of the offline phase

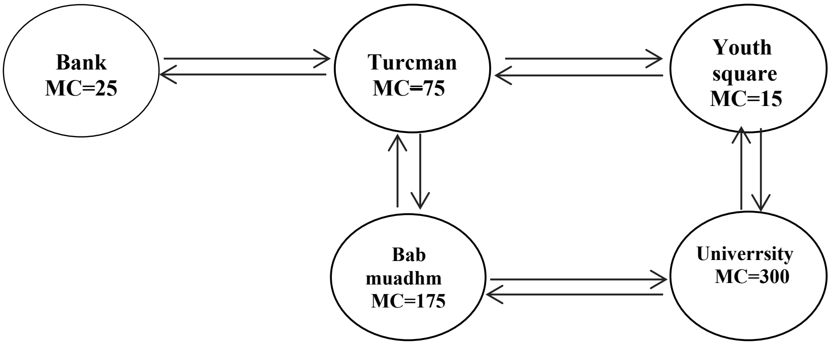


Fig. 4 Traffic jam graph of data clustering stage

trade-off between traffic jam and actual distance. The actual distance was calculated between road intersections; using the Euclidean distances, Fig. 7 shows the actual distance between each road intersections for the covered area in **Baghdad** city; the merging step is done using Eq. 2.

$$\mathrm{dis}_{\text{jam}-\text{dis}} = \mathrm{dis}_{\text{Actual}} + \eta_i * \frac{(\mathrm{MC}_{\text{ds}} - \mathrm{MC}_{\text{sr}})}{\mathrm{Vh}_{\text{c}}} \tag{2}$$

where $\mathrm{dis}_{\text{Actual}}$ is the distance between two adjacent road intersections and MC_{ds} and MC_{s} are the number of MC_{sr} (vehicles) in the destination and source road intersections, and adjusting factor ($\eta_i = 1$) is a factor that varies depending on the application. While Vh_{c} is a positive number, it indicates the number of vehicles that are equal to one kilometer (e.g., $\mathrm{Vh}_{\text{c}} = 7$ km/car means that the vehicle driver considers 1 km distance equivalent to the traffic jam of 7 cars). The following is a list of the relationships between the dist graph and the actual distance graph:

1. the distance in jam-distance graph is equal to the actual distance graph if the number of MC_{c} (vehicles) in the source and destination road intersections is equal due to $(\mathrm{MC}_{\text{ds}} - \mathrm{MC}_{\text{sr}}) = 0$, as illustrated in Fig. 5.
2. If the number of MC_{c} (vehicles) in the destination intersection is greater than the number of MC_{c} (vehicles) in the source intersection, that means, the vehicle

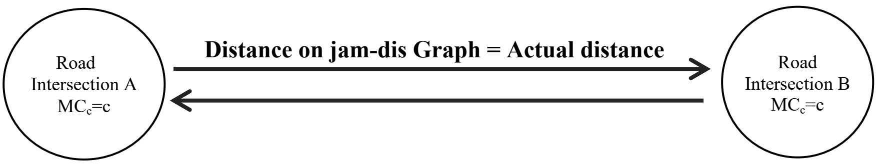


Fig. 5 Relationship between jam-distance graph and dist graph when MC is equal in source and destination road intersections

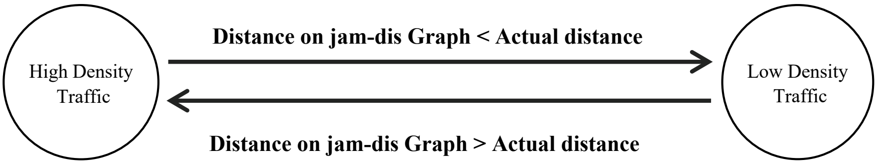


Fig. 6 Relationship between jam-distance graph and dist graph

moves from low-density traffic to high-density traffic, the distance in the jam-distance graph is larger than actual distance (not recommended path).

3. The distance in jam-distance graph is less than distance in actual graph. If the number of MC_c (vehicles) in the destination road intersection is less than the number of MC_c (vehicles) in the source intersection as shown in Fig. 6, the vehicle moves from high-density traffic to low-density traffic (Fig. 7).

Finally, the merging stage result is shown in Fig. 8. Illustrate the jam-distance graph.

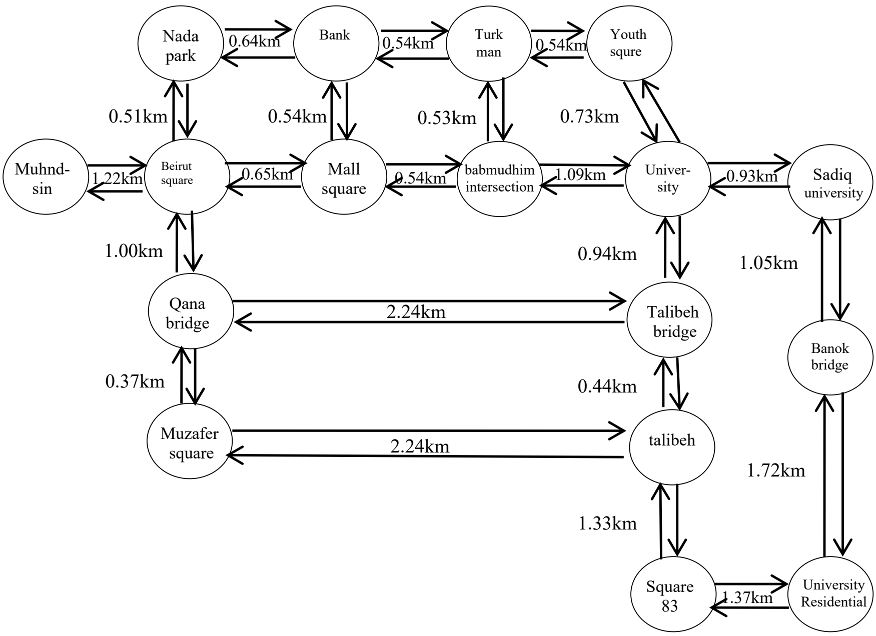


Fig. 7 Actual distance between each road intersection

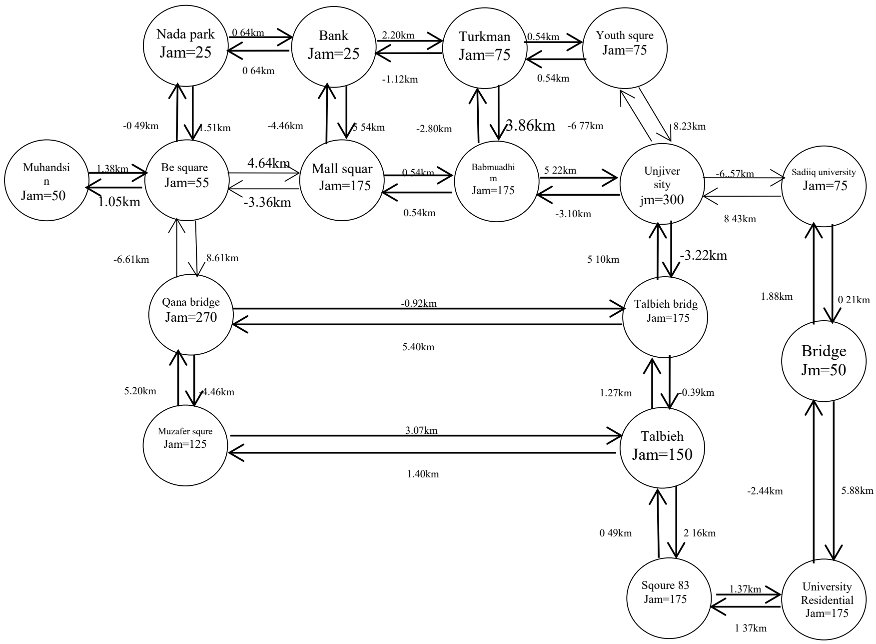


Fig. 8 Jam-distance graph for Baghdad city

3.3 Shortest Path Finding

This stage aims to find the shortest path in term of distance and traffic jam between vehicle driver (current location) and the destination; the Dijkstra's algorithm has been adopted to find the shortest path into the jam-distance graph.

4 Experiments Results

In order to evaluate the efficiency (running time) and the accuracy of the RN-CMS framework, it has been compared with Makariye [11] and Rachmawati [22]. The comparison is based on the dataset of traffic flow around AL-Mustansiriyah University in Baghdad city. The dataset represents the traffic flow of 10,000 vehicles from 6.00 am to 5:00 pm and covered about 100 km^2. The evaluation of the RN-CMS consists of two sections: In the first section, the parameters sensitivity for online and offline phase have been found which gives best performance. Secondly, the accuracy and the running time have been tested with the existing works as explained in the next section.

4.1 Parameter Sensitivity

The online phase supports the concepts of micro-clusters MC_s to summarize raw data, and the results of this phase significantly depend on the threshold value (dist.). The sum of square distance (SSQ) metric has been adopted to evaluate the online phase as defined in Eq. 3.

$$\mathrm{SSQ} = \sum_{u=1}^{Nx}\sum_{t=1}^{Nm}(x_t - \mu_{xu})^2 + (y_t - \mu_{yu})^2 \tag{3}$$

where $\boldsymbol{\mu}_{\mathrm{xu}}$, $\boldsymbol{\mu}_{\mathrm{yu}}$ represent the center of each micro-cluster in the covered area U, Nx is the total number of micro-cluster in the covered area, and Xt and Yt represent the vehicle coordinates. Nm represents the number of vehicle in the system. It is worth noting that the lower the SSQ value, the better the data summary, as the deviation of the micro-clusters $\mathrm{MCc_s}$ items from their center is lower. As a result, the typical dist value is (13 m), as indicated in Table 1.

In the offline phase, the parameters MinMccs and ε significantly affect the final result of the online results. Many experiments have been carried out using range of '$\boldsymbol{\varepsilon}$' values (0–0.008 km) and $\mathrm{minmc_cs}$ (1–10) to evaluate the performance of the offline phase based on (SSQ) the sum of square distance (SSQ). The small **SSQ** distance value means better compactness of each cluster; therefore, the value of $\mathrm{minmc_cs} = 4$ and $\varepsilon = 0.001$ refers much better clustering than other results as shown in Fig. 9.

Table 1 SSQ of distance threshold result

Distance	10	13	16	19	22
SSQ	15,523	14,432	15,432	16,343	17,564

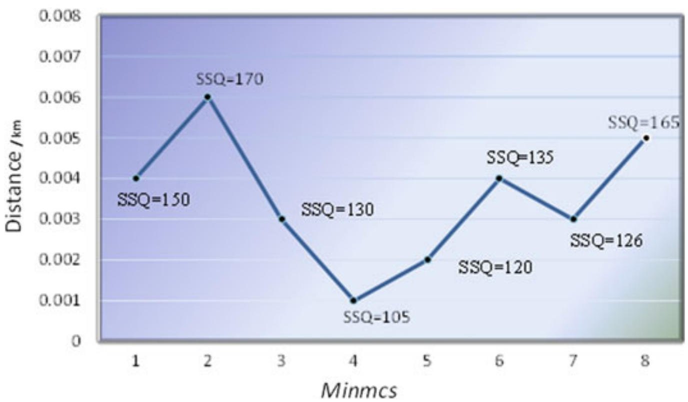


Fig. 9 Sum of squared distance analysis result

Table 2 RN-CMS accuracy result

	RN-CMS framework	Makariye [11]	Rachmawati [22]
Accuracy %	98	98	98

Table 3 RN-CMS efficiency result

	RN-CMS framework	Makariye [11]	Rachmawati [22]
Average running time for 100 inquires (s)	1.05	2.04	2.10

4.2 Accuracy and Running Time

In order to evaluate the accuracy and the efficiency (running time) of the RN-CMS framework, it has been compared with Makariye [11] and Rachmawati [22]; the three systems have been tested based on 100 inquires from various destinations and sources. The results illustrated that the RN-CMS framework has the highest accuracy as shown in Table 2.

Where the accuracy 99 means the proposed system has given the shortest path in term of traffic jam and distance. On the other hand, the efficiency (running time) of the three system has been shown in Table 3.

5 Conclusions

In this article, the RN-CMS framework is proposed to find the shortest path in road networks in terms of distance and traffic jam. The proposed framework consists of three stage. The first stage is to model a traffic jam graph through data summarization to immediately tackle the issue of high-speed GPS data generation and clustering the micro-cluster due to the massive size of raw data. While the second stage is the merging stage which constructs jam-distance graph based on actual distance and traffic jam finally, Dijkstra's algorithm has been used to find the shortest path according to the road traffic and actual distance. The experimental results show that the RN-CMS framework has similar accuracy (98%) as compared with Makariye N. (98%) and Rachmawati (98%). In the other hand, the running time of the RN-CMS is the shortest running time (1.05) s.

Acknowledgements The authors would like to expose their thanks and acknowledgments to Mustansiriyah University—Baghdad, Iraq, for supporting and encouragement. Thank you.

References

1. Mirge V, Verma K, Gupta S (2017) Dense traffic flow patterns mining in bi-directional road networks using density based trajectory clustering. Adv Data Anal Classif 11:547–561. https://doi.org/10.1145/2743025
2. Feng Z, Zhu Y (2016) A survey on trajectory data mining: techniques and applications. IEEE Access 4:2056–2067. https://doi.org/10.1109/ACCESS.2016.2553681
3. Bao J, Zheng Y, Wilkie D, Mokbel M (2015) Recommendations in location-based social networks: a survey. GeoInformatica 19:525–565. https://doi.org/10.1007/s10707-014-0220-8
4. Zheng Y (2015) Trajectory data mining: an overview. ACM Trans Intell Syst Technol 6:1–41. https://doi.org/10.1145/2743025
5. Riyadh M, Mustapha N, Sulaiman M, Sharef NBM (2017) CC_TRS: continuous clustering of trajectory stream data based on micro cluster life. Math Probl Eng 2017. https://doi.org/10.1155/2017/7523138
6. Galić Z (2016) Spatio-temporal data streams and big data paradigm. In: Spatio-temporal data streams. Springer, pp 47–69. https://doi.org/10.1007/978-1-4939-6575-5_3
7. Niu X, Chen T, Wu CQ, Niu J, Li Y (2019) Label-based trajectory clustering in complex road networks. IEEE Trans Intell Transp Syst. https://doi.org/10.1109/TITS.2019.2937882
8. Lytvynenko V, Lurie I, Krejci J, Voronenko M, Savina N, Taif MA (2019) Two step density-based object-inductive clustering algorithm. In: CEUR workshop proceedings, pp 117–135
9. Yuan G, Sun P, Zhao J, Li D, Wang C (2017) A review of moving object trajectory clustering algorithms. Artif Intell Rev 47:123–144. https://doi.org/10.1007/s10462-016-9477-7
10. da Silva TLC, Zeitouni K, de Macêdo JAF (2016) Online clustering of trajectory data stream. In: 2016 17th IEEE international conference on mobile data management (MDM). IEEE, pp 112–121. https://doi.org/10.1109/MDM.2016.28
11. Makariye N (2017) Towards shortest path computation using Dijkstra algorithm. In: 2017 international conference on IoT and application (ICIOT). IEEE, pp 1–3. https://doi.org/10.1109/ICIOTA.2017.807364
12. Shein TT, Puntheeranurak S (2018) Incremental clustering approach for evolving trajectory data stream. In: 2018 international electrical engineering congress (iEECON). IEEE, pp 1–4. https://doi.org/10.1109/IEECON.2018.8712334
13. Mao J, Song Q, Jin C, Zhang Z, Zhou A (2018) Online clustering of streaming trajectories. Front Comput Sci 12:245–263. https://doi.org/10.1007/s11704-017-6325-0
14. Shein TT, Puntheeranurak S, Imamura M (2018) Incremental discovery of crowd from evolving trajectory data. In: 2018 international conference on engineering, applied sciences, and technology (ICEAST). IEEE, pp 1–4. https://doi.org/10.1109/ICEAST.2018.8434397
15. Puntheeranurak S, Shein TT, Imamura M (2018) Efficient discovery of traveling companion from evolving trajectory data stream. In: 2018 IEEE 42nd annual computer software and applications conference (COMPSAC). IEEE, pp 448–453. https://doi.org/10.1109/COMPSAC.2018.00069
16. Zhang Y, Han LD, Kim H (2018) Dijkstra's-DBSCAN: fast, accurate, and routable density based clustering of traffic incidents on large road network. Transp Res Rec 2672:265–273. https://doi.org/10.1177/0361198118796071
17. da Silva TLC, Lettich F, de Macêdo JAF, Zeitouni K, Casanova MA (2020) Online clustering of trajectories in road networks. In: 2020 21st IEEE international conference on mobile data management (MDM). IEEE, pp 99–108. https://doi.org/10.1109/MDM48529.2020.00031
18. John A, Mathiyalagan P, Blessy A (2020) Continuous moving object clustering in dynamic road network. In: 2020 international conference on inventive computation technologies (ICICT). IEEE, pp 710–715. https://doi.org/10.1109/ICICT48043.2020.9112410
19. Wang J, Ye J, Yang Z. Traffic path planning algorithm based on shortest path algorithm
20. Li F, Shi W, Zhang H (2021) A two-phase clustering approach for urban hotspot detection with spatiotemporal and network constraints. IEEE J Sel Top Appl Earth Obs Remote Sens 14:3695–3705. https://doi.org/10.1109/JSTARS.2021.3068308

21. Li Y, Sun Q, Ji X, Xu L, Lu C, Zhao Y (2020) Defining the boundaries of urban built-up area based on taxi trajectories: a case study of Beijing. J Geovisualization Spat Anal 4:1–12. https://doi.org/10.1007/s41651-020-00047-6
22. Rachmawati D, Gustin L (2020) Analysis of Dijkstra's algorithm and A* algorithm in shortest path problem. J. Phys. Conf. Ser. 12061
23. Bansal R et al (2021) J Phys Conf Ser 1963:012170
24. Jeyalaksshmi S et al (2021) J Phys Conf Ser 1963:012145

Face Mask Waste Generation and Its Management During Covid-19

Anita Shrotriya, Pradeep Kumar Tiwari, Tarun Jain, Rishi Gupta, and Aditya Sinha

Abstract The pandemic during COVID-19 has had a negative influence on the world's fabric, including health systems, travel, living and working habits, and economies in numerous countries throughout the world. Furthermore, it has had a significant negative impact on continuing global attempts to curb excessive usage of plastic materials. The extensive usage by healthcare professionals and the overall community, of masks, sanitizers, and synthetic-based personal protective equipment (PPE) kits, has resulted in massive amounts of plastic trash, with no effective measures or policies in place to reduce its severity. Wearing a face mask as a way of protection against COVID-19 has become commonplace. However, because present mask disposal techniques (i.e., burning and reclamation) produce dangerous chemicals, huge production of contaminated face masks causes environmental difficulties. Furthermore, disposable masks are prepared of a variety of materials that are either non-recyclable or difficult to recycle. Therefore, as a result, it is critical to comprehend the scope of the problem and, equally essential, to devise a viable solution to contribute to the creation of a sustainable civic society.

Keywords COVID-19 · Face mask · PPE · Halyard H600 · N95

1 Introduction

The Waste increased in the spread of COVID-19 disease since March 11, 2020. It can be linked to an increase in healthcare wastes generated by medical amenities such as hospitals, quarantine centers, research laboratories, government and private clinics, and laboratories, in nearly every area of the world. WHO urged industry and government to raise production of surgical gloves, N95 masks, and standard surgical masks up to 40% to meet increased demand. As the demand of medical equipment increases,

A. Shrotriya · P. K. Tiwari (✉) · T. Jain · R. Gupta · A. Sinha
SCIT, Manipal University Jaipur, Rajasthan 303007, India
e-mail: pradeeptiwari.mca@gmail.com

A. Shrotriya
e-mail: Anita.shrotriya@jaipur.manipal.edu

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_41

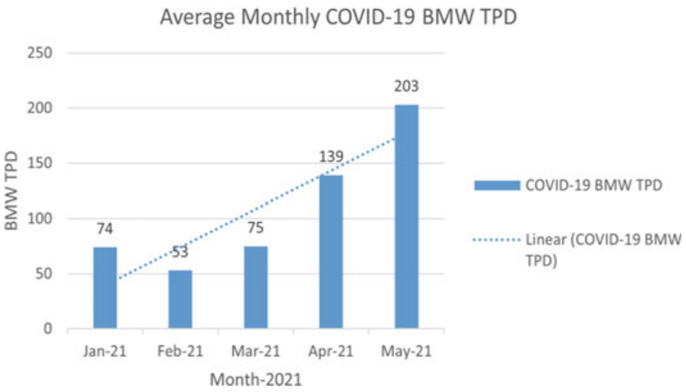


Fig. 1 COVID-19-related bio-medical waste production in India from Jan to May-21

India generated nearly 18,000 tons of COVID-19-related bio-medical waste between June and September, according to data provided by the Central Pollution Control Board (CPCB). Both healthcare personnel and patients employ personal protective equipment (PPE) such as gloves, face masks, head coverings, plastic coveralls, hazmat suits, and syringes. In recent years, there has been an increase in bio-medical waste due to COVID-19. In June, India produced 3025.41 tons of COVID-19-related bio-medical waste, which grew to 4253.46 tons in July; and in August and September, it reached 5238.45 tons and 5490 tons, respectively [1]. As said by Sourabh Manuja, Fellow, Center for Waste Management and Environment, the expansion in waste production is directly proportional to the increase in the number of instances.

Conferring to the WORLD HEALTH ORGANIZATION, healthcare waste comprises potentially hazardous bacteria that might affect other patients as well as healthcare personnel if not managed and disposed of appropriately, which includes COVID-19-related waste. So, the huge healthcare facilities' waste needs to be properly collected, segregated, disinfected, and finally disposed of to prevent the damaging impacts of both coronavirus and plastic wastes to avoid environmental pollution to a certain extent. In addition, if solid waste management is not deal properly, then it will be cause to the spread of the virus [2] (Fig.1; Table 1).

2 Materials and Chemical Components of Face Mask

Face masks provide user protection from a variety of external types of germs that may be present in the air they breathe. A good quality mask protects against germs, particles, particles of dust, body fluids. The mask materials and the chemical composition of various layers are chosen accordingly. The American Society of Testing Materials has specified the properties that a face mask should have five conditions that a good face mask should satisfy are—fluid resistance, particle filtration (PFE) efficiency,

Table 1 Top ten states/UTs of BMW generator in India (metric tons/day), Jan–May 2021

S. No.	Name of states/UTs	COVID-19 BMW (Tons/day)
1	Kerala	23.71
2	Gujarat	21.98
3	Maharashtra	19.02
4	Delhi	18.79
5	Karnataka	16.91
6	Uttar Pradesh	15.91
7	Tamil Nadu	13.57
8	Haryana	13.11
9	Andhra Pradesh	9.99
10	Madhya Pradesh	7.32

combustion, bacterial filtration efficiency (BFE), differential pressure. The standard N95 mask, found in India, has stretchable headbands to protect the mask on the user's face made of a thin layer of aluminum. The mask has four layers: an outside layer of spun-bond polypropylene, a second layer of cellulose/polyester, a third layer of molten polypropylene filter, and an inner (fourth) layer of polypropylene. To study the different structures of the different layers of the mask, these four layers are separated along with the metal strip and elastic cords. The material polypropylene, of which a large part of the mask is made, has good chemical resistance and can be processed through many modification methods such as injection molding and extraction [3]. Polypropylene is a polymer carefully prepared from propylene. Once the polymer is reused, its properties will be damaged. Cellulose is one of the most widely used polymers, waterless, and soluble in many solvents. Because the disposable mask is comprised of a variety of materials, it is difficult to reuse. As such, this work recommends a degree of environmentally friendly way, at the same time achieving the formation of maximum energy from the mask disposal. To date, thermo-compound interactions have been performed by CO_2. The key chemical components of the disposable mask were determined in the initial portion of this work: polypropylene, polyethylene, nylon, and aluminum. Asia, the most crowded landmass with billions of individuals living, will utilize a few large numbers of covers day by day. In contrast with careful face covers or N95 respirators, the hand-crafted face veils, i.e., cotton–wool covers show extremely restricted viability (greatest up to > 90% for particles > 300 nm). In this way, as the world's ability for assembling careful covers improves step by step, it will be prescribed to utilize careful face veils with airborne adequacy > 95% rather than cotton-based covers. This will lead to another test. Disposed of face veils will be seen wherever adjacent to shopping packs. In this way, while forcefully growing the ability to produce a great many facemasks every day, strong advances are needed to oblige the waste administration ramifications of countless careful veils and respirators.

3 Environmental Impacts of Face Masks

Face masks are currently a legitimate necessity in numerous public spaces around the planet. But, very little attention has been given on the most proficient method to discard or reuse them securely. Without better removal rehearses, an ecological debacle is approaching (Fig. 2).

Most of masks are fabricated from durable plastic-based materials, and whenever disposed of can persevere in the climate for quite a long time to many years. This implies they can affect the climate and individuals. This is genuinely dangerous on the grounds that despite the quick well-being risk they present, non-reusable covers are made from plastics like polypropylene which takes around 450 years to biodegrade [4]. With respect to the face masks which end up in the sea, they just add to the 8,000,000-ton mass of plastic entering the ocean consistently. The different plastics separate incredibly gradually into miniature plastics which at that point channel into marine natural pecking orders with unfortunate results. Miniature marine-plastics clutch poisons and toxins; plants and creatures at that point assimilate or ingest these substances, harming and executing them. It appears to be that face masks, fundamental as they are for forestalling the spread of coronavirus, are simply adding to effectively extraordinary worries over plastic contamination and the subsequent untamed life and human well-being impacts. Yet, face masks need not be added to this rundown of concerns; buying dispensable plastic face covers is profoundly superfluous given the thriving business sector of reusable ones, some of which utilize superb, socially expressive plans.

Regardless of whether purchasing or making a material face mask, a couple of things are sure: Moving away from expendable covers will set aside your cash and can create cash for a noble cause. It additionally energizes a social move for reusing

Fig. 2 Inappropriate disposal of mask during COVID-19 pandemic

a lot, fundamental for diminishing the interest for dispensable plastic face masks, and for lining up with more extensive disposition changes toward squander the board and opposing environmental change.

3.1 Risks to Individuals and Creatures

Initially, discarded masks may offer a risk of distributing COVID-19 to waste officials, garbage collectors, or the normal human being, public who have already gone over the garbage. Animals and plants are also altered in the intermediate to long term. Plastic garbage, by virtue of its sheer volume, has the ability to conceal conditions and separate living systems. A few species, too, cannot tell the difference between plastic and their feed, causing them to gag on rubbish [5].

Animals can grow starved even if they do not gag because the chemicals fill their stomachs but do not provide vitamins. More modest creatures may likewise get entrapped in the flexible inside the masks or inside gloves as they fall to pieces. Plastics degrade over time into smaller bits, and the longer rubbish is left in the environment, the more it degrades. Plastics break down first into microplastics, then into much more modest nano-plastics. These small particles and filaments are regularly enduring polymers that can gather in natural pecking orders. Only, one cover can create a great many particles, each with the possibility to likewise convey synthetics and microbes up the evolved way of life and conceivably even into people. Littered zones likewise will in general empower further littering, aggravating the issue.

3.2 Exposition

According to estimation calculated by researchers to combat COVID-19, World Health Organization projected that 89 million extra disposable masks were essential in medical and healthcare locations around the world. Similarly, Plastic Waste Innovation Hub at University College London published a new research, and according to that the current local interest in the UK is 24.7 billion veils in each year. However, if only reusable veils are used, the need for homegrown face covers in the UK falls significantly to roughly 136 million per year. Regardless, even with reusable covers, the design and method of cleaning have an impact. To estimate the whole environmental influence, the University College London group researched the production, use, and removal of disposable, recyclable, and reusable covers with expendable channels. After more than a year, they determined that machine washing reusable masks with no channels had the least effect. Hand washing masks increased the natural effect since, in comparison to machine washing, hand washing consumes more water and cleaner per cover. Expendable channels likewise increment the ecological effect in light of fact that the little channels are frequently produced using plastic like the dispensable masks, with a channel disposed of after each utilization [6, 7].

Maybe shockingly, the research paper determines that hand wash up reusable masks along with disposable channels had the greatest raised natural effect overall, even when compared to using entirely disposable covers. In view of the entirety of this, we should find a way to decrease the effect of wearing a face mask:

1. Use reusable covers that do not have expendable channels. Machine washes them on a regular basis, following the texture requirements.
2. Carry an extra mask with you so that if something goes wrong with the one you are wearing, you won't have to use or buy an expendable cover.
3. If you do have to utilize a dispensable cover, take it home (perhaps in a sack in the event that you need to take it off) and afterward set it on the right track into a canister with a top. If this is unimaginable, put it in a legitimate public receptacle.
4. Do not put expendable covers in the reusing. They can get trapped in expert reusing gear and be a likely biohazard to squander laborers.

3.3 Effective Disposal of Face Masks

In respect to waste management during COVID-19, most of the countries that produce extreme waste should be assess properly to manage medical waste disposals. It is also required to take important and vital steps to prevent the possible disperse of communicable and infectious viruses via masks, and their dumping processes and healthcare waste like masks must be handled by ignition and safe landfill. As a result, we can say that to dispose of COVID-19-correlated medical waste, high-temperature ignition takes precedence over disposal of such medical equipment and masks. This is the utmost popular, naturally secure, and appropriate approach for destroying remnants of the virus using a superior-level boiler temperature, with the high temperature and time set to 1100 °C for three minutes. During COVID-19, in China, for example, emergency medical waste ignition includes the use of municipal solid waste incinerators to co-process medical waste for disposal in rotary furnaces with temperatures greater than 850 °C. In some areas where incinerators are unavailable, before being disposed of in sanitary landfills, medical waste can be sterilized at high temperatures and boiled [8].

Nonetheless, during the COVID-19 epidemic, the entire capacity for medical waste disposal is woefully inadequate. Landfills, for example, remain a popular MSW management option in developing Asian countries such as India, Bangladesh, Thailand, Myanmar, and Malaysia. As a result, acute processes such as discrimination, sorting, storage, assortment, transference, and last removal are mandatory for the supervision of used mask and medical waste. As a result, COVID-19's proliferation in the environment and throughout numerous countries should be restricted. In India, used face masks are kept in a paper bag for 72 h before being discarded as usual waste in confined or other residences. To avoid reuse, face mask straps are removed before disposal. Greater Chennai Corporation (GCC) has advised non-quarantined families and residents to disinfect worn masks with a 5% bleach solution or a sodium hypochlorite solution before discarding them [9]. Before being distributed to sanitary staff, masks are draped and stored in a locked waste bin. GCC treats them as home hazardous garbage and incinerates it [10].

4 Conclusion

Wearing a mask is the minimum one can do for protection against coronavirus. Mask works best when everyone wears it and maintain social distancing. Now, the question arises which type of masks is more supportable? In terms of sustainability or supportability, we look up for reusable masks rather than buying disposable masks. Due to sudden large-scale production of masks, PPE kits, face shields, etc., a huge amount of medical waste is getting produced. A major part of this waste is plastic based, due to which it creates a huge burden on the environment. Although the material of the masks is considered bio-degradable, but in today's scenario, the decomposition of such huge quantities of waste is not effectively taking place. One of the ways to tackle this problem is the use of reusable home-made cloth masks. Home-made masks are not considered as medical masks which means they may not have the required pore sizes that offer security against the coronavirus.

These types of masks, according to studies, are also said to be affective against other viruses, dust particles, etc. The University of Florida developed a novel mask out of Halyard H600 medical cloth, which is also used to wrap sterile gowns, towels, tool kits, and procedure trays. The University of Florida created a unique mask out of Halyard H600 medical cloth, which is also used to wrap up hygienic gowns, towels, tool kits, and procedure trays. These sterilization covers have a microbiological barrier, and the masks passed a qualitative suitability test on many people with a sugary or sour taste. It may be more effective than a surgical mask, but it should not be used in place of N95. The mask was invented by Sonia Mehta, a physician, and her colleagues. She described how she began with an idea that grew into a full-scale manufacture of nearly 5000 masks in three weeks, and how she discovered ways to minimize, reuse, and recycle N95 masks.

References

1. Saadat S, Rawtani D, Hussain CM (2020) Environmental perspective of COVID-19. Sci Total Environ 728:138870
2. Peng J, Wu X, Wang R, Li C, Zhang Q, Wei D (2020) Medical waste management practice during the 2019–2020 novel coronavirus pandemic: experience in a general hospital. Am J Infect Contr 48(8):918–921
3. World Health Organization. Water, sanitation, hygiene, and waste management for the COVID-19 virus: interim guidance

4. Net XH (2020) Medical waste treatment in Wuhan from emergency to stability. www.xinhuanet.com/local/2020-04/01/c_1125796126.htm
5. Asian Development Bank (2020) Managing infectious medical waste during the COVID-19 pandemic
6. You S, Sonne C, Ok YS (2020) COVID-19's unsustainable waste management. Science 368.6498(2020):1438; Minoglou M, Gerassimidou S, Komilis D (2018) Healthcare waste generation worldwide and its dependence on socio-economic and environmental factors. Sustainability 9.2(2017):220
7. Korkut EN (2018) Estimations and analysis of medical waste amounts in the city of Istanbul and proposing a new approach for the estimation of future medical waste amounts. Waste Manage 81(2018):168–176
8. Rahman MM, Bodrud-Doza M, Griffiths MD, Mamun MA (2020) Biomedical waste amid COVID-19: perspectives from Bangladesh. Lancet Global Health 2020:8
9. Abu-Qdais HA, Al-Ghazo MA, Al-Ghazo EM (2020) Statistical analysis and characteristics of hospital medical waste under novel Coronavirus outbreak. Glob J Environ Sci Manage 6(4):1–10
10. Kulkarni BN, Anantharama V (2020) Repercussions of COVID-19 pandemic on municipal solid waste management: challenges and opportunities. Sci Total Environ 743(2020):140693

Design of Wireless Sensor Network (WSN) for Healthcare Application

Noora Kamil Flayyih and Ibrahim A. Murdas

Abstract Development of a method is based on wireless sensor for determining pollutant material in water and agriculture products. The identification is based on a smart algorithm. Pollutant material detection by sensors has many properties such as being easy to achieve and very cheap. And at the same time, the detecting mechanism gives result approach to real analysis with mini difference in concentration. This detection method describes the analysis of different pollutant materials that affect human health and may cause health risk such as cancer disease. The suggested method depends on three different types of sensors to collect the data. The principle work of the system is based on the attitude and behavior of ions at different temperatures; the collected data processed using microcontroller to get digital signal and the reading have been dealing with it to get a comparison between control sample (free sample) and measured sample. A smart system is used to process the data for training the system to distinguish between two pollutant materials by using advanced algorithm. The detecting pollutant materials reach 98% at different concentrations.

Keywords Pollutant materials · Sensors · Light absorbance · Smart algorithm

1 Introduction

The general concept of pollutant materials from the aspect of food safety and agriculture organization [1] is a widely distributed material, or compositions of materials that have been detected in soil, water, precipitate, an aquatic biota, and municipal sewage can be processed to obtain a semiliquid waste usually used as a fertilizer [2, 3], and pollutant materials is aiming optimized solutions and rationalized consumption of environmental resources so it is important repelling, avoiding, controlling, or eliminating any pest or harmful weeds if safety of commodities and consumers from

N. Kamil Flayyih (✉) · I. A. Murdas
University of Babylon, Babil/Hilla, Iraq
e-mail: noorakamil87@gmail.com

I. A. Murdas
e-mail: eng.ibraheem.abdullaa@uobabylon.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_42

side effects of pollutant materials is being accounted as in environmental effect so the term pollutant materials are indispensable to modernistic agriculture for protecting crops whether pre- or post-harvest stages [4] involve substances meant to use as insect management or regulators for plant growth by disease prevention. Many vegetables and fruits are contaminated with pollutant materials and quantities of hazardous chemicals enter the chain of food which is being caused by the overuse of pesticides and the intensification of these habit in agriculture is caused due to an accumulation of pesticide residues in an environment and is remained for long periods and that seriously is posing a risk on the environment and ecosystem of worldwide and people in contact with pesticides with various exposure routes by ingestion, skin, respiratory tract [5] which have adverse effects on human health with infection in many diseases of psychiatric [6, 7], endocrine disorders [8], disrupting renal, neurological problems, hepatic and reproductive turbulence in addition to its effect in plants and animals even at low concentration [9–11]; especially, there is utterly consciousness in health aspects in society being happened recently [12–14] and for protecting consumers strict legislation has been in place and the agricultural commodities must be suitable for international regulations [15]. Organophosphate and Pyrethroid are one class of pollutant materials which are widely used in agriculture. Pollutant materials have been classified according to use as fungicide, nematicide, molluscicide, and rodenticides [16–18], and concerning the chemical composition commonly can be classified as organochlorine, pyrethroids, organophosphates, and carbamate which are widely used because of their high activity and comparatively low immutability in the environment [19]. In the short term, exposure to a large number of chemical quantities during short periods is being caused relatively acute toxicity while in long term, exposure to a little number of chemical quantities during long periods is being caused relatively chronic toxicity.

For the guarantee of food quality and safety, analytical instruments are required for easy and fast quantification and monitoring of pollutant material residue and can be used at low concentrations [6, 20, 21]; it is extremely necessary for assurance compliance with legal levels [22–24], depending on praxis of pollutant materials and to action alongside detection ways have been confirmatory, so detection system is being designed where it is characterized to get best performance and low cost where it can be used by different category consumer which has vulnerable intake, and short period results with fast analysis, without laborious treatment to sample and permits direct detection and reinforces toxicological risk assessment process, in addition, the possibility of system miniaturization to be easily carried or transport which make easy field trials and no need to be highly trained with the capability of real-sample monitoring in situ without laboratory, the system designed depends on ions and light and its behavior in aqueous solution, according to all above, that make the detection device is a premium elect for development a portable detection system. The aim of the study was to develop a method for this purpose and several sample analyses were performed to examine whether altering various samples would cause improvement of the reference method which had been used and after that acquired data can be trained by neural network [25] to identify the type of pesticide residue by classifying acquired data into two or more class according to (0, 1) classifying and also unknown pesticide residue values can be excellent predicted especially if the output of the neural network is equal to the target value where that make the error equal zero and the network training is perfect to predict unknown value and make identification to know the type of pollutant materials if belong to type a or B or more than one type; the proposed system will be explained further [26].

2 Working Principles and Device Designing

2.1 The Principle of Work

Depending on the attitude of ions and light in an aqueous solution and how the ions move between the electrode of the sensor and the electrolyte.

TDS principle work: it consists of two parts; first part is the TDS meter and the other is the probe which has two electrodes separated by a space when the probe is impressed in an aqueous solution that contains dissolved solids, the dissolved solids generate ions then and voltage source will be applied; then, there will be the transportation of ions which generated by dissolved solid and electrons which generated by voltage source between two electrodes that led to generating an electrical conductivity and we can measure it by TDS meter as shown in Fig. 1.

PH principle work: ph. sensor also consists of two parts; the first part is the ph. meter consists of a BNC connector, Trimmer for calibrating, and ph. sensor interface circuit pinout, and the other part is the probe consists of reference electrodes and ph. electrode which measures the concentration of H^+ ions in the solution, PH generated by the solution caused by the difference in electrical potential between ph. electrode and reference electrode and electrical potential are generated and it can be measured by ph. meter (ph. sensor interface circuit pinout) as shown in Fig. 2.

The principle work to the electrode can be explained in the form of electrical admittance which is being founded between two electrodes that have been put in an electrolyte as shown in Fig. 3:

C_d = capacitance (double-layer) of the electrode/electrolyte interface.

G_b = the conductance of bulk electrolyte.

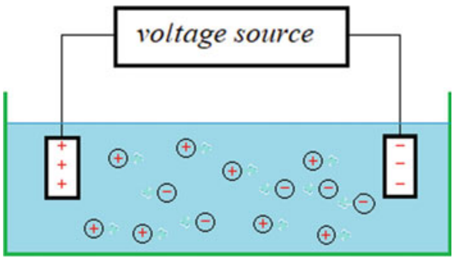


Fig. 1 Shows TDS principle work

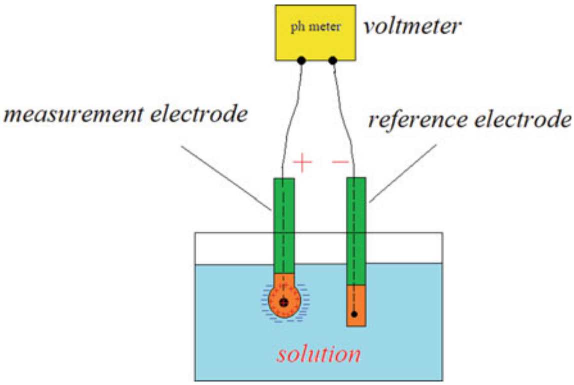


Fig. 2 Shows PH principle work

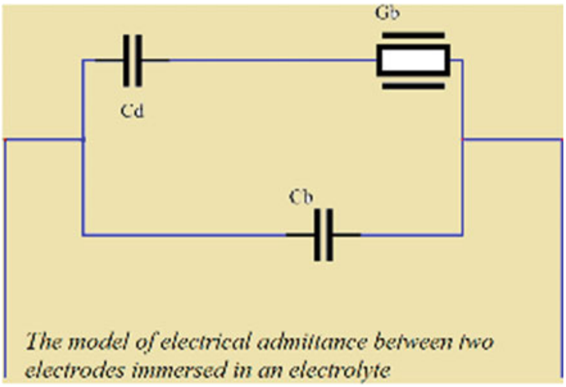


Fig. 3 The model of electrical admittance between two electrodes

C_b = the two electrodes geometrical capacitance which apart by the measured electrolyte.

The conductance changes when if frequency is 100 kHz or less and it is being constant if the frequency is above this value.

At frequency above 100 kHz, there is an increase in conductance being affected by concentration not on the frequency the conductivity is increasing by increasing the number of ions in electrolyte per unit volume at dilation an increase in concentration will cause an increase in volume and increasing in volume of solution will cause a decrease in conductivity as the relationship conductivity = (number of ions/volume) number of ions per unit volume; this is called Molar conductance (C_m); at lower frequencies, it will decrease at increased frequency and it will be constant or saturated at high frequencies, where (C_d) acts as a short circuit at high frequencies and ($G_m = G_b$).

2.2 Device Designing

The work includes two stages; the first stage is design system detection of pollutant materials residue and the other is to make identification to the data detection by using a smart system.

First stage: the system detection consists of three parts; the first part is the TDS sensor which depends on electrical conductivity, and the second part is the ph. sensor which depends on electrical potential, and the last part is LDR which depends on light absorbance and temperature had been controlled by using a temperature sensor. After the designing was being completed, a code to each part must be written to operate each sensor part and then each part of the device system detection must be calibrated and then the data which had been collected will be classified into two or more types by using identification by neural network, and also it can be predicted with the use of perceptron network and fuzzy logic; this is representing the second stage we can classify the data to show this data for which class of pesticide belongs. This classification can be done by separating this data with a linear line. This line can be controlled by changing the value of bias and to get the best performance the error must be equal zero; we can get that if (a is equal to t) where (a) is the output of the perceptron network and t which represented the correct target that required so the difference between target and the output is equal the error, if error equal zero then the output equal the target and the performance to the network will be the best and if a is equal to zero and t is equal to 1; therefore, the error is the result of subtracting the output a from the target *t*. So the error is equal to one, and the input vector will classify one or if the output a is equal to one, and the target t is equal to zero; therefore, the error is the result of subtracting the output a from the target *t*. So the error is equal to -1, and the input vector will classify as zero. We can adjust the Wight until we get the required values which approach from three cases if the transfer function is hardlim then the output will be only either 0 or 1. We can summarize the whole process as follows:

1. Design the detection system device.
2. plot the circuit diagram.
3. write sensors code Arduino IDE software version 1.6.13.
4. Calibration to each sensor and adjust the temperature at the range of 25 °C.
5. take a result.
6. Make identification by classifying data that is taken from sensors entering the data as input and target vectors and creating perceptron network by function (newp) after that we must sure that the data can be separated by perceptron network.
7. The data for two classes can be separated by a constant K which represented any number by which the two data types can be parted to avoid interference between data.
8. initiate weight and bias.
9. training the network by function train (net, *p, t*).
10. make simulation

$$a = \mathrm{sim}(\mathrm{net},\, p) \tag{1}$$

11. determine the output.

$$a = \mathrm{hardlim}(W(1)p2 + b(1)) \tag{2}$$

12. Comparing the output with the target vectors of the network for getting the error.

$$e = t - a \tag{3}$$

13. using function learn p to adjust the weight and bias to get a new weight and bias repeating the training process until the error becomes zero or as minimum error as possible.

$$W^{\mathrm{new}} = W^{\mathrm{old}} + ep^{T} \tag{4}$$

$$b^{\mathrm{new}} = b^{\mathrm{old}} + e \tag{5}$$

$$w = \mathrm{net}.iw\{1,\, 1\} \tag{6}$$

$$b = \mathrm{net}.b\{1\} \tag{7}$$

$$W = W + dw \tag{8}$$

All presses above had been expressed by a block diagram as shown in Fig. 4.

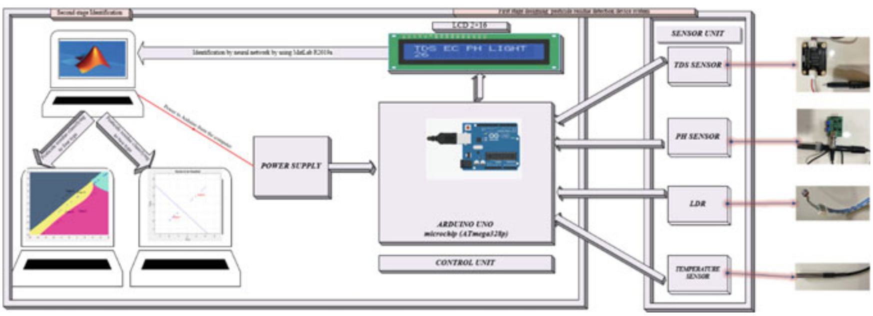


Fig. 4 Block diagram of pesticide residue detection system device

3 Data Process

The data which had been trained by using a smart system-based neural network consists of set of data to train the network. Four types of pollutant material residue are the first Bif 181 samples, Bifpure 9 samples, GlyMi 181 samples, and Glypure 7 samples; those four types of files contain data to four types of pollutant materials by making simulation of the data by MATLAB R2019a software. The four types had been classified by using perceptron code and we can predict unknown value by entering the higher value and the lowest value and the value between them can be predicted to which class it belongs. We can do that also by using fuzzy logic. We made three codes: the first one to classification BifMI and GlyMi, the second to classification Glypure and Bifpure, and the third code combines the four types and the data had been classified into four classes which belonged to four types of pollutant materials. All the proposed work and methodology can be explained further with a flowchart in Fig. 5.

4 Experimental Results and Discussion

4.1 Results of a Detection System

Results: The analysis of pollutant materials for Gly. (48%SL) shows that there is a pollutant material residue in three types of sample water and orange juice. We can distinguish the value of pollutant material residue incrementing by comparing the difference in values between the free pollutant materials (control) market sample value and containing pollutant materials market sample. So the difference between those values is representing there is a ratio in pollutant materials residue in samples. We can notice from charts that all the sensor parts indicate changes in their value from the control sample except EC. It cannot detect pollutant material residue sample only, where it remains constant and is good in the water and orange juice; in addition to the other sample, water and orange juice is good in detecting and also LDR is good in detecting water and orange juice but it is different in the magnitude of pollutant material residue detecting in the sensor part and the response rate for sensing the pollutant materials residue in samples, as shown in (Figs. 6, 7, 8 and Tables 1, 2).

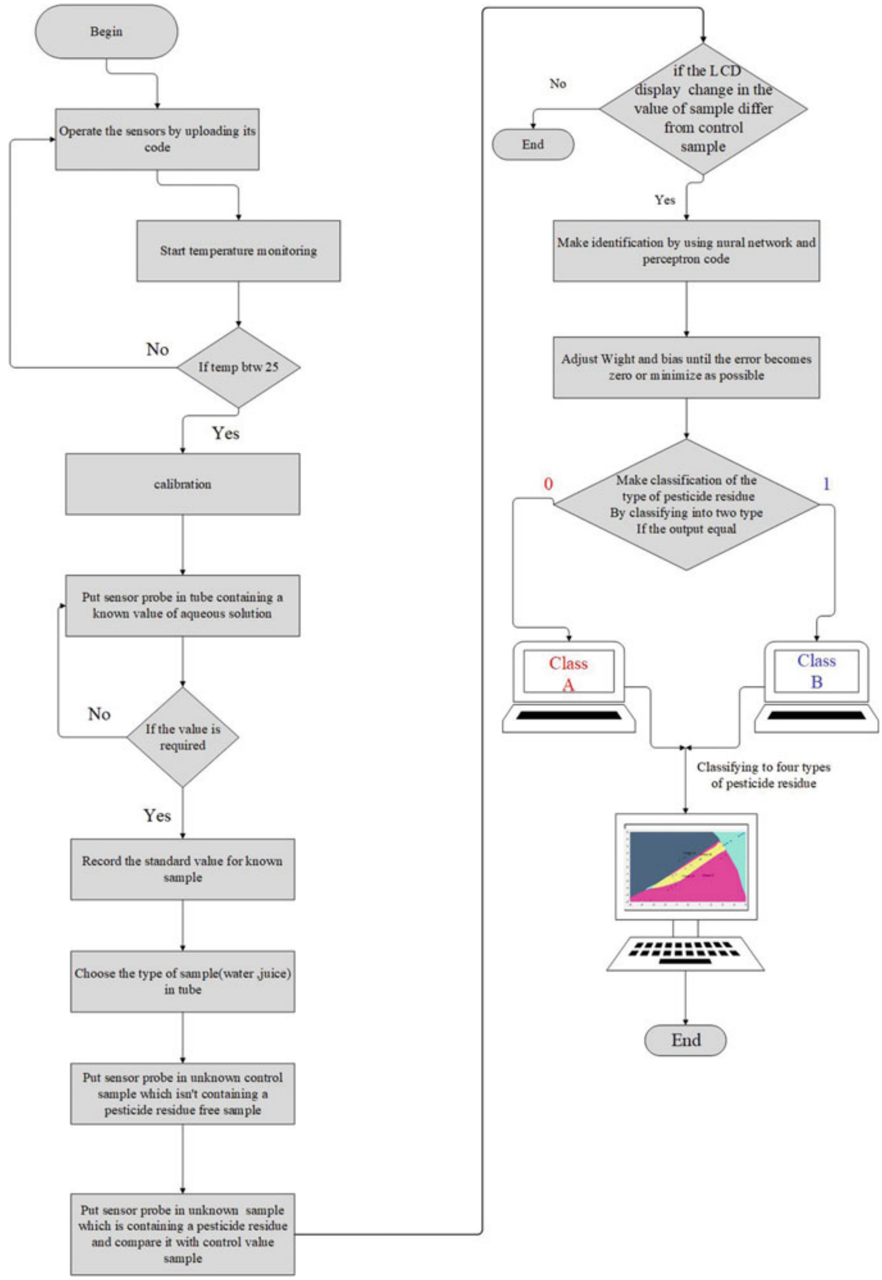


Fig. 5 Flowchart of the proposed work or methodology

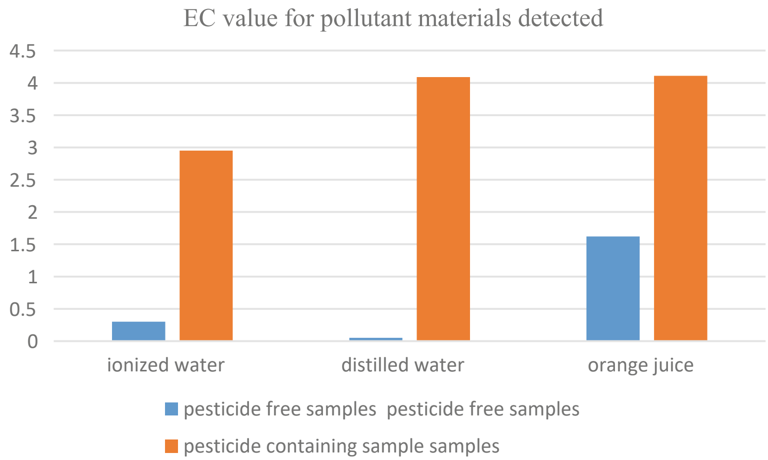


Fig. 6 Chart indicates the EC value of the different concentrations of market sample

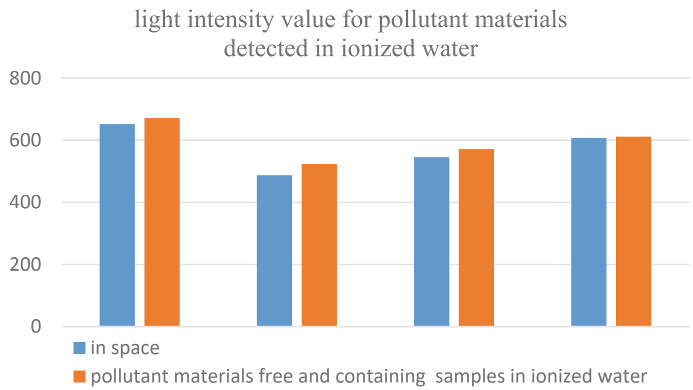


Fig. 7 Chart indicates the light intensity value for ionized water of the different concentrations of market samples

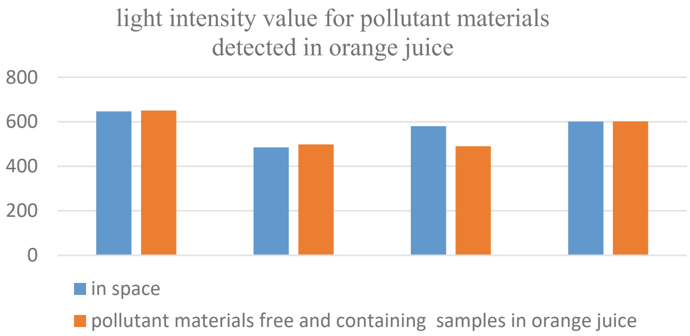


Fig. 8 Chart indicates the light intensity value for the orange juice of the different concentrations of market samples

Table 1 The results of the statistical calculations to conductivity EC values

The analyzed type	The number		Arithmetic mean to the EC values		Standard deviation	Relative % deviation
	Free market samples	Containing market samples	Pollutant materials free market samples	Pollutant materials containing market samples	Pollutant materials containing samples	Pollutant materials containing samples
Ionized water	1	11	0.3	2.95	1.782	60.320
Distilled water	1	5	0.05	4.09	0	0
Orange juice	1	5	1.62	4.11	0.24	5.8495

Table 2 The results of the statistical calculations to light intensity values

The analyzed type	The number	Arithmetic mean to the light intensity values		Standard deviation	Relative% deviation
	Free and containing market samples	In space	Pollutant materials containing market samples	Pollutant materials containing samples	Pollutant materials containing samples
Ionized water	5	652	671.6	6.269	0.9334
		487	524	18.88	3.6
		545	570.6	14.53616	2.547
		608	611.6	10.526	1.721
Orange juice	5	647	650.8	7.9498	1.22
		485	498.4	15.789	3.168
		580	589.8	17.838	3.024
		601	601.8	6.5345	1.0858

5 Conclusion

A detection method is proposed to monitor pollutant materials in water, and market fruit juice samples depending on the characteristics of different types of sensors and take several functions to overcome various types of vegetable and fruit and low concentration detecting with speed in response and good sensitivity at low concentration, so we had taken different types of sensors. The sensor depends on the principle of electrical conductivity, three types of sensor that mean more properties and more benefit to overcome more sample and more low concentration. The aim of this system is to detect any change in the reading of sample compared with the control sample. After testing the sample, the results indicate a change in reading when adding different amount of pollutant materials, and the results had been used successfully in detecting pollutant materials. The results are used to detect pollutant material residue in water and fruit juice. In this study, we took orange juice, and the other parameter PH also is good in testing water and juice so if one parameter is not good in detecting system, it can be compensated in properties of other parameters; by this way, we can overcome more samples.

References

1. World Health Organization (2016) Manual on development and Use of FAO and WHO specifications for pesticides, 3rd Revision; World Health Organization, Geneva
2. Domínguez I, González RR, Liébanas FJA, Vidal JLM, Frenich AG (2016) Automated and semi-automated extraction methods for gc–ms determination of pesticides in environmental samples. Trends Environ Anal Chem 12:1–12
3. Cheng Z, Dong F, Xu J, Liu X, Wu X, Chen Z, Pan X, Zheng Y (2016) Atmospheric pressure gas chromatography quadrupole-time-of-flight mass spectrometry for simultaneous determination of fifteen organochlorine pesticides in soil and water. J Chromatogr 1435:115–124
4. Hassani S, Momtaz S, Vakhshiteh F, Maghsoudi AS, Ganjali MR, Norouzi P, Abdollahi M (2017) Biosensors and their applications in the detection of organophosphorus pesticides in the environment. Arch Toxicol 91:109–130
5. Verma N, Bhardwaj A (2015) Biosensor technology for pesticides—a review. Appl Biochem Biotechnol 175:3093–3119
6. Mostafalou S, Abdollahi M (2013) Pesticides, and human chronic diseases: evidence, mechanisms, and perspectives. Toxicol Appl Pharmacol 268:157–177
7. Harrison V, Ross SM (2016) Anxiety and depression following cumulative low-level exposure to organophosphate pesticides. Environ Res 151:528–536
8. Campos É, Freire C (2016) Exposure to non-persistent pesticides and thyroid function: A systematic review of epidemiological evidence. Int J Hyg Environ Health 219:481–497
9. Kumar J, Melo JS (2017) Overview on biosensors for detection of organophosphate pesticides. Curr Trends Biomed Eng Biosci 5:555–663
10. Kazemi M, Tahmasbi AM, Valizadeh R, Naserian AA, Soni A (2012) Organophosphate pesticides: a general review. Agric Sci Res J 2:512–522
11. Jaga K, Dharmani C (2006) Ocular toxicity from pesticide exposure: a recent review. Environ. Health Prev Med 11:102–107
12. Selvi S, Karthikeyan R, Vanitha U (2012) Organic farming: Technology for environment-friendly agriculture. In: Proceedings international conference on advances in engineering, science and management (ICAESM), pp 132–136

13. Damalas CA, Eleftherohorinos IG (2011) Pesticide exposure, safety issues, and risk assessment indicators. Int J Environ Res Publ Health 8:1402–1419
14. Aktar W, Sengupta D, Chowdhury A (2009) Impact of pesticides use in agriculture: their benefits and hazards. Interdiscip Toxicol 2:1–12
15. Capoferri D et al (2018) Affinity sensing strategies for the detection of pesticides in food. Foods 7(9):148
16. Garcia FP, Ascencio SYC, Oyarzun JCG, Hernandez AC, Alavarado PV (2012) Pesticides: classification, uses and toxicity. Measures of exposure and genotoxic risks. J Res Environ Sci Toxicol 1:279–293
17. Wong A, Silva TA, Caetano FR, Bergamini MF, Marcolino-Junior LH, Fatibello-Filho O, Janegtiz BC (2017) An overview of pesticide monitoring at environmental samples using carbon nanotubes-based electrochemical sensors, C, 3, 8
18. Mulchandani A, Chen W, Mulchandani P, Wang J, Rogers KR (2001) Biosensors for direct determination of organophosphate pesticides. Biosens Bioelectron 16(6):225–230
19. Scognamiglio V, Arduini F, Palleschi G, Rea G (2014) Biosensing technology for sustainable food safety. TrAC Trends Anal Chem 62:1–10
20. Rotariu L, Lagarde F, Jaffrezic-Renault N, Bala C (2016) Electrochemical biosensors for fast detection of food contaminants—trends and perspective. TrAC Trends Anal Chem 79:80–87
21. Della Pelle F, Vázquez L, Del Carlo M, Sergi M, Compagnone D, Escarpa A (2016) Press-printed conductive carbon black nanoparticle films for molecular detection at the microscale. Chem Eur J 22:12761–12766
22. Kumar P, Kim K-H, Deep A (2015) Recent advancements in sensing techniques based on functional materials for organophosphate pesticides. Biosens Bioelectron 70:469–481
23. Eddleston M, Buckley NA, Eyer P, Dawson AH (2008) Management of acute organophosphorus pesticide poisoning. Lancet 371:597–607
24. Chowdhary S, Bhattacharyya R, Banerjee D (2014) Acute organophosphorus poisoning. Clin Chim Acta 431:66–76
25. Changlong W et al (2009) Rapid detection of pesticide residues in vegetables with mini-colorimeter. In: 2009 9th international conference on electronic measurement & instruments. IEEE
26. Najm H, Hoomod HK, Hassan R (2020) A proposed hybrid cryptography algorithm based on GOST and salsa (20). Period Eng Natural Sci 8(3):1829–1835

Optimal Control for Robot–Environment Interaction in Robotic Systems

Dao Phuong Nam, Nguyen Trung Nghia, Bui Thi Hai Linh, and Nguyen Hong Quang

Abstract The reference generator in the robot control system depends on unknown robot–environment interaction. In this paper, an adaptive reinforcement learning (ARL)-based optimal control is developed to achieve the solution, which considers an unknown environment as a linear discrete-time system. A performance index that establishes the interaction effects of trajectory tracking error and external torque is optimized using admittance adaptation. The proposed reference generator is designed in a completed control system of robot systems under unknown robot–environment interaction. Simulation studies are conducted to show the effectiveness of the proposed solution.

Keywords Robot–environment interaction · Adaptive reinforcement learning (ARL) · Admittance adaptation · Unknown environment

1 Introduction

Robot–environment interaction with holonomic constraints broadly exists in robotic systems and is one of the most significant influences to lead the complexity for designing the control system [1]. Until recently, plenty of control methods for robotic systems have been mentioned, for example, sliding mode control in Bilateral Teleoperators [1], the backstepping technique in mobile robotic systems [2–4], and optimal control in mobile robots [5–7]. Generally, the control strategies can be categorized into two major classes, including nonlinear control approaches and optimal control methods. However, almost control approaches for robotics have not considered the influence of robot–environment interaction [1–7] and have focused on the nonlinear control method. The optimality principle approach for finding controllers in robotics has been discussed in [5–10] with the obstacle of solving Hamilton–Jacobi–Bellman

D. P. Nam · N. T. Nghia
Hanoi University of Science and Technology, Hanoi, Vietnam

B. T. H. Linh · N. H. Quang (✉)
Thai Nguyen University of Technology, Thai Nguyen, Vietnam
e-mail: quang.nguyenhong@tnut.edu.vn

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_43

and Ricatti equations. The key idea to overcome these disadvantages is the implementation of the Adaptive Reinforcement Learning (ARL) technique [11–16] with several main approaches, such as Q learning [11], discrete-time control [12–16], and data-driven [16]. The ARL technique in optimal control enables us to find the solution based on an iterative algorithm via the convergence of control policy and Bellman function. It should be noted that the solutions in [8, 10] were investigated by solving Ricatti equation in continuous-time systems. However, this paper proposed the discrete-time control-based ARL to generate the reference trajectory from the desired trajectory and external torque. The proposed algorithm can obtain an appropriate reference trajectory under unknown robot–environment interaction in the robot control system.

2 Preliminaries and Problem Statement

In this section, we investigate a manipulator interacting with the external environment, as shown in Fig. 1 and the following equation:

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + g(q) = \tau - \tau_{\text{ext}} \tag{1}$$

where $M(q), C(q, \dot{q}), g(q)$ are the system matrix, coupling matrix, and gravity matrix, respectively. $\tau \in \Re^n$ is the control input vector. $\tau_{ext} \in \Re^n$ is defined as a joint torque vector to be modeled unknown environment as:

$$\alpha\dot{q} + \beta q = -\tau_{\text{ext}} \tag{2}$$

where α, β are unknown damping and stiffness matrices, respectively. It can be seen that $\tau_{\text{ext}} \in \Re^n$ generates the effect to robot systems via unknown environment model (2). Due to the interaction between end-effector and environment to be described by

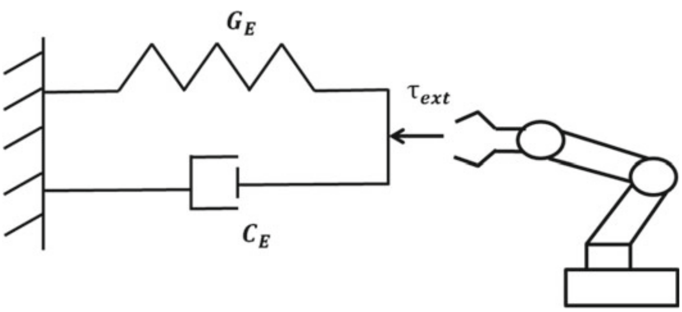


Fig. 1 Model of interacting region [8]

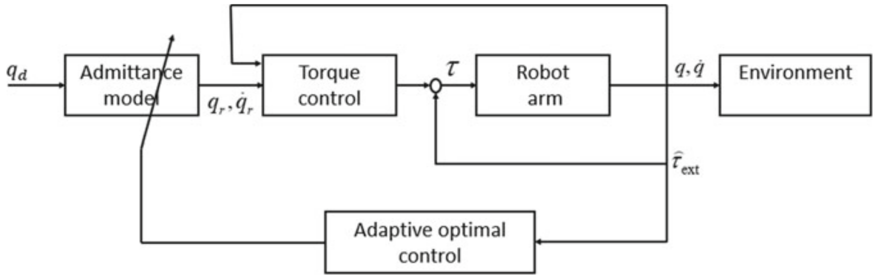


Fig. 2 The control system for manipulators with robot–environment interaction [8]

$\tau_{\text{ext}} \in \Re^n$, the reference in control system (Fig. 2) is modified from desired trajectory q_{d} to reference trajectory q_{r}, $\frac{\text{d}}{\text{d}t}q_{\text{r}}$ to guarantee the tracking problem of joint variable.

The control objective is to obtain the torque $\tau_{\text{ext}} \in \Re^n$ in Eq. 2 satisfying the optimal interaction performance with defined cost function:

$$J = \int_0^\infty ((q - q_{\text{d}})^T Q(q - q_{\text{d}}) + \tau_{\text{ext}}^T R\tau_{\text{ext}})\text{d}\tau \tag{3}$$

By finding $\tau_{\text{ext}} \in \Re^n$ to guarantee the minimum of cost function (3), the desired interaction is obtained with the advantage of efficiently achieving q_{r}, $\frac{\text{d}}{\text{d}t}q_{\text{r}}$ as discussed in next sections.

3 Optimal Control for Generating the Reference Trajectory

In this section, we focus on establishing the reference q_{r}, $\frac{\text{d}}{\text{d}t}q_{\text{r}}$ in control system (Fig. 2) using optimal control design. First, discrete-time adaptive reinforcement learning (ARL) is presented for the unknown model. Second, a reference trajectory generator is proposed with ARL algorithm.

3.1 Discrete-Time Adaptive Reinforcement Learning

Consider a discrete-time linear model obtaining from continuous-time system by discretizing model:

$$\xi_{k+1} = A_{\text{d}}\xi_k + B_{\text{d}}u_k;\ y_k = C\xi_k \tag{4}$$

The control objective is to find control input obtaining the minimum of following performance index:

$$V = \sum_{i=k}^{\infty} (y_i^T Q_{\mathrm{d}} y_i + u_i^T R_{\mathrm{d}} u_i) \tag{5}$$

Based on the work in [9], we proposed the modified off-policy algorithm using damping coefficient $0 < \gamma < 1$ to improve the convergence velocity as follows:

Algorithm 1: Discrete-Time ARL Algorithm

Choose an initial control signal $\mu_k^0 = \mu^0$. Then, for $j = 0, 1, \ldots$, implement until convergence.

Step 1 Updating the Value with damping coefficient $0 < \gamma < 1$:

$$\overline{z}_{k-1,k-N}^T \overline{P}^{j+1} \overline{z}_{k-1,k-N} = y_k^T Q_{\mathrm{d}} y_k + (u_k^j)^T R_{\mathrm{d}} u_k^j + \gamma \overline{z}_{k,k-N+1}^T \overline{P}^j \overline{z}_{k,k-N+1} \tag{6}$$

Step 2 Partitioning matrix $\overline{P}$ as in $\overline{P} = \begin{bmatrix} p_0 & p_u & p_y \\ p_u^T & P_{22} & P_{23} \\ p_y^T & P_{32} & P_{33} \end{bmatrix}$. The control Policy is updated as:

$$u_k^{j+1} = \mu^{j+1}(\xi_k) = -(R_d/\gamma + p_0^{j+1})^{-1} (p_u^{j+1} \overline{u}_{k-1,k-N+1} + p_y^{j+1} \overline{y}_{k,k-N+1}) \tag{7}$$

3.2 *Admittance Adaptation for Robot–Environment Interaction*

According to (2), we can extend the state space representation with the additional term q_{d} as:

$$\dot{\xi} = \begin{bmatrix} -\alpha^{-1}\beta & 0 \\ 0 & K \end{bmatrix} \xi + \begin{bmatrix} -\alpha^{-1} \\ 0 \end{bmatrix} \tau_{\mathrm{ext}};\ \dot{\xi} = A\xi + B\tau_{\mathrm{ext}};\ \dot{q}_{\mathrm{d}} = K q_{\mathrm{d}} \tag{8}$$

Subjected to the performance index:

$$V = \int_0^{\infty} \left((q - q_{\mathrm{d}})^T Q (q - q_{\mathrm{d}}) + \tau_{\mathrm{ext}}^T R \tau_{\mathrm{ext}}\right) \mathrm{d}\tau = \int_0^{\infty} \left(\xi^T Q' \xi + \tau_{\mathrm{ext}}^T R \tau_{\mathrm{ext}}\right) \mathrm{d}\tau \tag{9}$$

After discretizing system (8), we develop Algorithm 1 for obtaining Admittance Adaptation in the control system (2).

Step 1 Using the control policy

$$u_k = -(R_{\mathrm{d}}/\gamma + p_0^0)^{-1}(p_u^0 \overline{u}_{k-1,k-N+1} + p_y^0 \overline{y}_{k,k-N+1}) + e_k \tag{10}$$

with e_k to be considered the additional term for satisfying PE condition.

Step 2 Using data collection u_k and y_k to find the matrix $\overline{P}$ with the condition for stopping the Algorithm:

$$\left|\overline{P}^{j+1} - \overline{P}^{j}\right| < \varepsilon \tag{11}$$

Step 3 Based on optimal control:

$$\tau_{\mathrm{ext}} = -K\xi = -R^{-1}B^T P * \xi = f(q_{\mathrm{r}}, q_{\mathrm{d}}) \tag{12}$$

We achieve the reference trajectory in control system (Fig. 2).

4 Simulation Results

The proposed algorithm is evaluated for the planar robot with the following parameter table by MATLAB-Simulink software (Table 1).

The control system in Fig. 2 is established by the frame of simulink and m.file code to verify the admittance adaptation law with ADP technique. The actual environment model is described: $0.01\dot{q} + (q - 0.3) = -\tau_{\mathrm{ext}}$, and the desired trajectory is chosen:$q_{\mathrm{d}} = \begin{bmatrix} 0.5 - 0.2e^{-t} \\ 0.3 - 0.3e^{-t} \end{bmatrix}$. It should be noted that, under proposed algorithms 1 and 2, the response of performance index is shown in Fig. 3. Additionally, the optimal matrix is computed with the following terms:

$$p_0^* = 0.1671 \times 10^{-3},\ p_{u1}^* = 0.1312 \times 10^{-3};$$
$$p_{y0}^* = \begin{bmatrix} -0.016 & -0.0664 \end{bmatrix};\ p_{y1}^* = \begin{bmatrix} -0.0039 & -0.0664 \end{bmatrix}$$

Table 1 Parameter of planar robot

m_1	Weight of joint 1	2.0 kg
m_2	Weight of joint 2	2.0 kg
l_1	Length of link 1	0.2 m
l_2	Length of link 2	0.2 m
I_1	Inertia moment of joint 1	0.027 kg m^2
I_2	Inertia moment of joint 2	0.027 kg m^2
g	Gravity	9.8 m/s^2

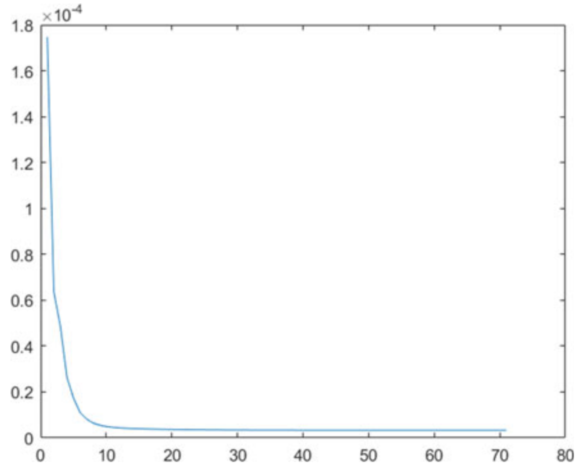


Fig. 3 The convergence of performance index

Therefore, the impedance control is achieved from algorithms 1 and 2 as follows:

$$
\begin{aligned}
q_{rk} = \frac{1}{0.0125}(u_k &+ 0.1026u_{k-1} - 0.0519(q_{dk} - 0.3) \\
&- 0.0031(q_{dk-1} - 0.3) - 0.0519(q_{dk-1} - 0.3))
\end{aligned}
$$

5 Conclusions

In this paper, the reference trajectory is generated by a modified discrete-time ARL algorithm with a damping coefficient. The algorithm is implemented under off-policy framework. Additionally, the impedance adaptation is combined with an inner control loop in a completed control system. Finally, the convergence of performance index in simulation results verified the effectiveness of the proposed algorithm in that the planar robot is operated under the influence of robot–environment interaction.

Acknowledgements This research was supported by the Thai Nguyen University of Technology.

References

1. Liu Y, Dao PN, Zhao KY (2020) On robust control of nonlinear teleoperators under dynamic uncertainties with variable time delays and without relative velocity. IEEE Trans Industr Inf 16(2):1272–1280. https://doi.org/10.1109/TII.2019.2950617
2. Dao PN, Nguyen HQ, Ngo MD, Ahn SJ (2020) On stability of perturbed nonlinear switched systems with adaptive reinforcement learning. Energies 13(19):5069

3. Binh NT, Tung NA, Nam DP, Quang NH (2019) An adaptive backstepping trajectory tracking control of a tractor trailer wheeled mobile robot. Int J Control Autom Syst 17(2):465–473
4. Nguyena T, Hoang T, Pham M, Dao N (2019) A Gaussian wavelet network-based robust adaptive tracking controller for a wheeled mobile robot with unknown wheel slips. Int J Control 92(11):2681–2692
5. Dao PN, Liu YC (2020) Adaptive reinforcement learning strategy with sliding mode control for unknown and disturbed wheeled inverted pendulum. Int J Control Autom Syst:1–12
6. Dao PN, Loc PT, Huy TQ (2021). Sliding variable-based online adaptive reinforcement learning of uncertain/disturbed nonlinear mechanical systems. J Control Autom Electr Syst:1–10
7. Dao PN, Nguyen HQ, Nguyen TL, Mai XS (2021) Finite horizon robust nonlinear model predictive control for wheeled mobile robots. Math Probl Eng
8. Yang C, Peng G, Li Y, Cui R, Cheng L, Li Z (2018) Neural networks enhanced adaptive admittance control of optimized robot–environment interaction. IEEE Trans Cybern 49(7):2568–2579
9. Liu X, Ge SS, Zhao F, Mei X (2020) Optimized impedance adaptation of robot manipulator interacting with unknown environment. IEEE Trans Control Syst Technol 29(1):411–419
10. He W, Xue C, Yu X, Li Z, Yang C (2020) Admittance-based controller design for physical human–robot interaction in the constrained task space. IEEE Trans Autom Sci Eng 17(4):1937–1949
11. Kiumarsi B, Lewis FL, Modares H, Karimpour A, Naghibi-Sistani MB (2014) Reinforcement Q-learning for optimal tracking control of linear discrete-time systems with unknown dynamics. Automatica 50(4):1167–1175
12. Liu Y, Zhang H, Luo Y, Han J (2016) ADP based optimal tracking control for a class of linear discrete-time system with multiple delays. J Franklin Inst 353(9):2117–2136
13. Luo B, Yang Y, Wu HN, Huang T (2019) Balancing value iteration and policy iteration for discrete-time control. IEEE Trans Syst Man Cybern Syst
14. Wei Q, Li B, Song R (2017) Discrete-time stable generalized self-learning optimal control with approximation errors. IEEE Trans Neural Netw Learn Syst 29(4):1226–1238
15. Wei Q, Lewis FL, Liu D, Song R, Lin H (2016) Discrete-time local value iteration adaptive dynamic programming: convergence analysis. IEEE Trans Syst Man Cybern Syst 48(6):875–891
16. Luo B, Liu D, Wu HN (2017) Adaptive constrained optimal control design for data-based nonlinear discrete-time systems with critic-only structure. IEEE Trans Neural Netw Learn Syst 29(6):2099–2111

Improving the Ability of Persons Identification in a Video Files Based on Hybrid Intelligence Techniques

Lubna Thanoon ALkahla and Jamal Salahaldeen Alneamy

Abstract With the advancement in technology, the importance of person recognition in photographs or videos has grown due to its usefulness in the search for wanted persons and criminal identification using various theories and algorithms and various non-hybrid and hybrid techniques to identify the person through the face, and its features have been developed. The paper proposed a hybrid algorithm to improve the performance of person identification in video files. The system works through several steps: first, face detection using Viola–Jones algorithm; second, feature extraction by algorithm local binary pattern (LBP); final, person identification by hybrid proposed algorithm (HPBFF) by hybrid between backpropagation neural network and firefly algorithm. The results show that the system was able to identify and monitor the person with a high classification accuracy rate of 98.4%, compared to 94.7% for the approach without the hybrid. The results of the tests revealed that the system is robust and has a high recognition rate, making it suitable for use in mobile and compact identification and authentication.

Keywords Face recognition · Firefly · LBP · Video detection · Backpropagation neural network · Viola–Jones method

L. Thanoon ALkahla (✉)
Department of Computer Science, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq
e-mail: lubna.csp21@student.uomosul.edu.iq; lubna.thanoon@uoninevah.edu.iq

L. Thanoon ALkahla · J. Salahaldeen Alneamy
Department of Software, College of Information Technology, Ninevah University, Mosul, Iraq
e-mail: jamal_alneamy@uomosul.edu.iq; jamal_alneamy@uoninevah.edu.iq

J. Salahaldeen Alneamy
Department of Software, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_44

1 Introduction

In the fields of computer vision, protection, biometrics, and surveillance, the development of effective and reliable person recognition systems is a major research focus. Person identification is usually accomplished by observing and measuring biologic characteristics, or biometrics, where a biometric is a distinguishing attribute that is used for recognition. For example, identification of the people that have similar faces, fingerprints, or iris [1].

Recently, the number of video cameras has increased in different public places, and the quality of the data that obtained from a standard surveillance camera has been improved and converted to digital format. And in the same time, the average cost of receiving, processing, and transitioning the data has decreased. These reasons led to the creating and development of the methods and algorithms for the effective identification of people from photograph images and video, and the most of these algorithms and methods focusing on the recognition of human faces [2].

Face recognition is a technique of recognition that is used to find the faces of people whose images were saved in the dataset. Face recognition has always been a common topic of research compared with the traditional card recognition, iris or fingerprint because it is people's facile method of personal identification and because of its non-meddling nature [3].

Face recognition is one field of biometric security. This technology is used for a variety of purposes such as in a law enforcement. This technology is increasing in most countries. Police collect shots from criminals and compare them with face recognition databases [8]. Once a criminal's photograph has been taken, their images will be added to databases to be scanned whenever police carry out another criminal search. [4]

Artificial intelligence (AI) is one of the most interesting and universal areas of computer science. AI is now everywhere around us, and it has ushered in a new era in the world by allowing intelligent machines to exist. It is currently working with a variety of subfields, and the machine learning (ML) is a subfield of an artificial intelligence (AI) that developed involving algorithms for self-learning that derive information from data to make predictions. Instead of requiring humans, manually extract rules and construct models from massive amounts of data [5].

The amazing advancements in the field of artificial intelligence techniques and machine learning have enabled the development of algorithms that have superior capabilities to identify people through their face recognition techniques. Researchers have suggested several different models of artificial neural networks as a result of the success of artificial neural network algorithms in the field of recognition and identification of people. The challenge is to determine the most suitable model that can be effectively utilize in this field [6].

All presented identification algorithms and methods have their own advantages and disadvantages related to the method of obtaining human metrics (needed physical contact with a person or not needed), reliability of results, and cost [7]. Therefore, the principle of hybridization was adopted in order to obtain the advantages of all the methods that were used and to build accurate and effective systems and models at the same time that were easy to apply [8].

This paper is structured as follows: Sect. 2: describes the proposed system and introduces algorithms used and suggested algorithms that build of this work, Sect. 3: datasets that were used for evaluated the proposed system, Sect. 4: The results are presented and discussed, and Sect. 5: conclusions, final section represents the approved references.

2 Proposed Methodology

In this proposed system, we will focus on identifying people from the video sequence based on faces. The proposed mechanism in this system is to detect and track persons with a high rate of accuracy at the same time with high process speed. This proposed system searches for persons within a video file and tracks them, where each video file is divided into a group of frames, each frame consists of a face or a number of faces, where all faces are identified in one frame by the algorithm of face detection, we use Viola–Jones. The features are then extracted for these faces through feature extraction algorithm, and we use local binary pattern (LBP). We have a matrix made up of a bunch of rows, and each row represents the features of a specific face. To improve the system's performance, hybrid artificial intelligence technology is used to identify person by hybrid backpropagation and firefly techniques (HBPFF).

3 Preprocessing

We preprocess the images in the database and the images of the wanted persons. We used special treatment, and first, it normalized all images with size (120 * 120), and second, enhancement of images by adjustment the contrast and brightness by using histogram equalization (HE).

4 Face Detection

Face detection is the first step in face recognition systems [9], which is a computer technology to define the face for the persons in the images [10], and it is characterized as the process of recognizing faces in a visual scene [11]. The aim is extracting the facial area from an image. It also has several applications such as retrieving a specific image based on content and monitoring a large group of people in addition to smart computer interfaces. Many techniques are used to detect faces. The most important algorithm used is Viola–Jones [8, 12].

4.1 Viola–Jones Algorithm

Paul Viola and Michael Jones discovered one of the most important methods of identifying a face in 2001. The following characteristics of the Viola–Jones algorithm make it an effective detection algorithm:

(1) Robust—high detection rate, (2) face detection only, not recognition—distinguishing faces from non-faces, and (3) real time—at least two frames per second for realistic applications [13].

The major components of this algorithm are [13, 14]

1. Haar feature selection

   All human faces have certain characteristics in common. Haar features should be used to match these regularities (see Fig. 1). The area of the nose bridge is lighter than the eyes, which is a typical feature of human faces [14], and the area of the eyes is darker than the area around the upper cheeks.

2. Creating an integral image: This is an image that tests rectangular features.
3. AdaBoost training: AdaBoost training is a learning algorithm for choosing the right features and training classifiers that use them.
4. Cascading classifiers: are a series of levels, each of which contains a powerful classifier, and all features are divided into stages, each with a certain number of features.
5. Decide if a given feature is obviously not a face or may be a face at each stage.

5 Feature Extraction

The feature extraction expresses to extract of features from the data available in a database that is used from the beginning so that they are used in machine learning algorithms for classification. The aim of this process is to use low-level features to identify each image in the database [8]. These features, known as descriptors, were previously used to determine image similarity. As compared to the original image, descriptors are usually smaller [15]. One of the most widely used methods for extracting features is the LBP algorithm.

Fig. 1 Haar features

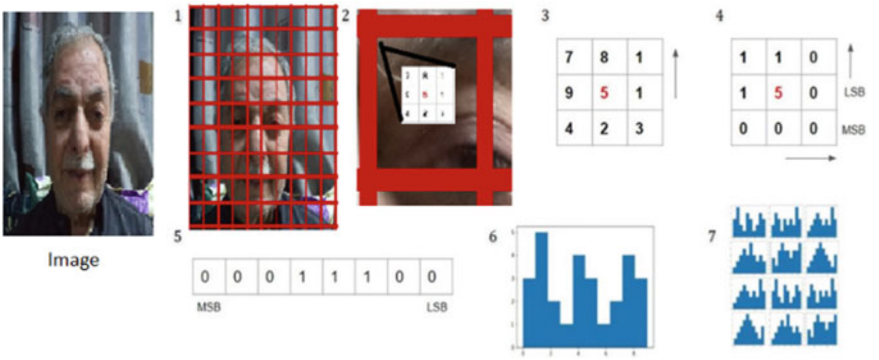


Fig. 2 Steps of the LBP algorithm

5.1 *Local Binary Pattern (LBP)*

Ojala et al. [16] introduced LBP as an important descriptor. This method is widely used in image processing applications. The LBP code of a center pixel is created by encoding the computed threshold value into a decimal value, and the LBP code of a neighboring pixel is generated by encoding the computed threshold value into a decimal value. The mathematical description of LBP is as follows:

$$\text{LBP} = \sum_{j=0}^{P-1} \text{S}(n_i - c_c)^{2i} \tag{1}$$

$$s(x) = \begin{cases} 1, if x > 0 \\ 0, otherwise \end{cases}$$

P stands for the number of neighboring pixels, n_i stands for the ith neighboring pixel, and c_c stands for the center pixel. The histogram features of size $2P$ are extracted using the obtained LBP code. As a consequence, the vector length of the histogram function for six adjacent pixels is 64 (Fig. 2).

5.2 *Identification*

Identification is the process of comparing an unknown face to a huge database of known faces in order to determine the identity of the unknown person. Identification can be employed on either “cooperative” or “uncooperative” people who are aware that they are being scanned. At the definition level, the data was trained and then used to identify whether the persons in the video are or are not [17], and at this stage, the data was trained and selected in two ways. Normal backpropagation neural network (NBPNN) and HBPFF (proposed algorithms).

5.3 Normal Backpropagation Neural Network (NBPNN)

Backpropagation is a common way of training artificial neural networks to reduce the objective function to the smallest possible value. It is a generalization of the delta rule and is a supervised learning approach. It necessitates a training set that contains a dataset of the desired output for a variety of inputs. BP network is a multilayer feedforward network trained by the error backpropagation algorithm. Its structure is similar to the human brain and is composed of a large number of neurons connected to each other. Data is multiplied by weights and moved from the input layer to each nerve in the next layer. The data from the previous layer is passed by adding and transferring functions to achieve the output of each nerve in the mid-layer. Because of the total amount of data coming from the mid-layer, an evaluation procedure is carried out in the output layer by repeating the processes in the previous layer [18]. The mean squared error (MSE) between the actual and expected outputs is computed as a result of this calculation by this equation:

$$MSE = \frac{1}{2}\sum_m (B_m - \varsigma_m^{\varsigma})^2$$

where B_m is expected output, ς_m^{ς} is the actual output. The training will be done using the error backpropagation process, which involves updating each weight on the network. It is attempted to obtain the least possible error, namely the optimum value, by repeating this procedure [19].

5.4 HBPFF Algorithm

The proposed algorithm combines the firefly algorithm and the back propagation technique. A starting population is generated first in the algorithm, as illustrated in Fig. 3. The number of fireflies in the population must be determined while constructing the starting population.

The fact that individuals are chosen may make the algorithm run faster, but it also means that the algorithm examines solution sets with fewer fireflies. As a result, the possibilities of discovering the best option will be reduced. Having a larger number of fireflies will improve the probability of discovering the best option. Excessive numbers raise computation expenses and can generate a workload that today's computers cannot handle [20].

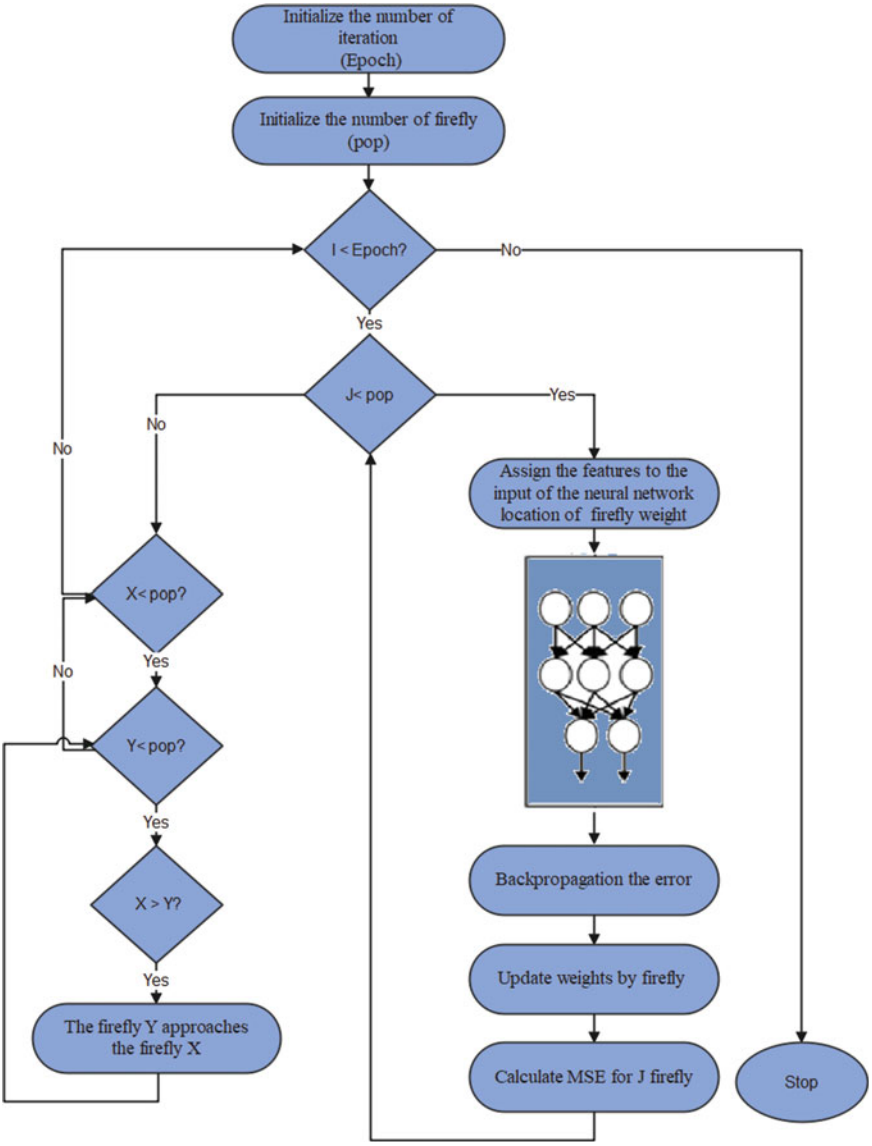


Fig. 3 Flowchart of proposed algorithm (HBPFF)

The algorithm defines a starting location to the fireflies when the parameters are determined. The I loop allows the algorithm to be run as many times as the step number specified. Every time this loop repeats itself. For each firefly, it activates the j loop. The completeness of the artificial neural network is modified as a result of the output and error value it obtains, while the backpropagation algorithm changes the position of the fireflies. Following the end of the j loop, the x and y loops run in a concentric way to allow the fireflies to interact with one another. For each firefly, the x and y loops are run. As a result, the light density of one firefly may be compared to that of another. The purpose function's brightness is inversely proportional to the brightness of the firefly. The MSE value acquired from each firefly is computed as the brightness of the firefly in an inversely proportional way since the objective function inside the algorithm is to reduce the error of the artificial neural network.

6 Result

In this research, a group of video files was adopted to identify and search for persons within these video files. A database of 400 images of people's faces in different situations was relied upon. For the training process, the proposed algorithm was applied. In the process of training, testing and comparing the results with a normal backpropagation neural network.

When the proposed hybrid algorithm was applied to the databases during the training phase, the results were obtained: 98.4% accuracy was obtained, and the MSE was 0.1512, while when relying on the normal neural network, 94.7% classification accuracy was obtained, and the MSE was 0.2333.

Table 1 and Fig. 4 show the percentage of accuracy and MSE when using the two algorithms HBPPSO and NBPNN; with feature extraction algorithm, LBP shows how much accuracy and less MSE have been improved using a proposed algorithm.

Finally, the HBPFF and NBPNN algorithms were tested on a video. It was identified for three persons inside a video file to find out whether these persons are present in this video file or not, and Table 2 and Fig. 5 show the persons who were identified

Table 1 Percentage of accuracy and MSE of HBPPSO and NBPNN algorithms

Algorithms	Accuracy (%)	MSE
HBPFF	98.4	0.1512
NBPNN	94.7	0.4622

Fig. 4 Shows the percentage of accuracy and MSE for HBPPSO and NBPNN algorithms

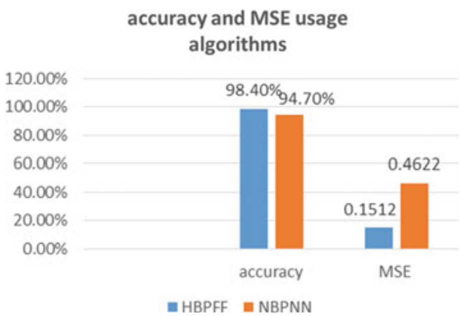


Table 2 Shows result use proposed hybridization algorithm and the normal neural network

Video	NBPNN	HBPFF
P1	1	1
P2	1	1
P3	0	1
Identification rate	66.66%	100%

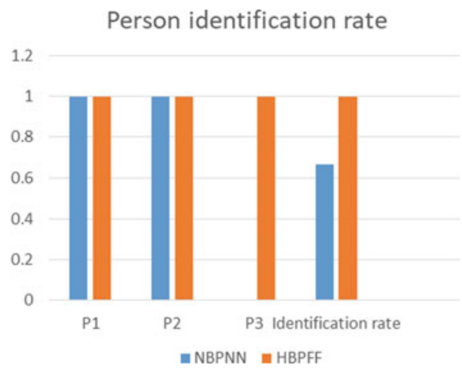


Fig. 5 Shows the result of person identify in a video

based on the proposed hybridization algorithm and the identification rate compared to the normal neural network.

7 Conclusion

In this paper, a proposed hybridization algorithm was proposed, which is (HBPFF) to identify person in a video file, where the nature-inspired firefly algorithm is merged with backpropagation for optimization in neural network training. The results showed that the proposed algorithm gave high classification accuracy and low mean square error compared to the normal backpropagation neural network algorithm without hybridization, as well as identification rate. The number of persons inside the video file, based on the proposed algorithm, was higher compared to the number of persons who were identified based on the normal method. The results of the tests revealed that the system is robust and has a high recognition rate, making it suitable for use in mobile and compact identification and authentication.

References

1. Anil K (2011) Introduction of biometric integrated automated fingerprint identification system, (IAFIS) http://www.fbi.gov/abouus/cjis/fingerprints_biometrics/iafis/iafis
2. Richa V, Kaushik B (2019) A survey on various face detecting and tracking techniques in video sequences. In: International conference on intelligent computing and control systems, (ICCS). https://doi.org/10.1109/iccs45141.2019.9065483
3. Masupha L, Zuva T, Ngwira S, Esan O (2015) Face recognition techniques, their advantages, disadvantages and performance evaluation. In: International conference on computing, communication and security (ICCCS), pp 1–5. https://doi.org/10.1109/CCCS.2015.7374154
4. Amazing uses for face recognition–facial recognition use cases (2020), https://www.facefirst.com/amazing-uses-for-face-recognition-facial-recognition-use-cases/
5. Bahrammirzaee A (2010) A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. Neural Comput Appl 19(8):1165–1195. https://doi.org/10.1007/s00521-010-0362-z
6. Le TH (2011) Applying artificial neural networks for face recognition. Adv Artif Neural Syst article id 673016 (2018). https://doi.org/10.1155/2011/673016
7. Wang K et al (2018) Survey on person re-identification based on deep learning. CAAI Trans Intel Technol 3(4). https://doi.org/10.1049/trit.2018.1001
8. Alkahla L, Alneam JS (2021) Face identification in a video file based on hybrid intelligence technique-review. J Phys Conf Ser 1818(1):012041. 1https://doi.org/10.1088/1742-6596/1818/1/012041
9. Shankerpandey U, Chhabra S (2016) Asurvey on face detection methods and feature extration of face recognition using pca. Indian J Comput Sci Eng (IJCSE)
10. Lewis,Michael B, Ellis HD (2003) How we detect a face: a survey of psychological evidence. Int J Imag Syst Technol 13:3–7
11. Kumar A, Kaur A, Kumar M (2019) Face detection techniques: a review. Artif Intel Rev 52(2):927–948. https://doi.org/10.1007/s10462-018-9650-2
12. Wang Y (2014) An analysis of the viola-jones face detection algorithm. Image Process On Line on 2014–06–26. ISSN 2105–1232 c IPOL, https://doi.org/10.5201/ipol.2014.104
13. Dabhi MK, Pancholi BK (2016) Face detection system based on viola-jones algorithm. Int J Sci Res (IJSR) 5(4):62–64. https://doi.org/10.21275/v5i4.nov162465
14. Haar-like feature (2020), https://en.wikipedia.org/wiki/haar-like_feature
15. Kamal S, Jalal A (2016) A hybrid feature extraction approach for human detection, tracking and activity recognition using depth sensors. Arab J Sci Eng 41:1043–1051
16. Rodriguez Y (2006) Face detection and verification using local binary patterns. A PH.D. thesis no 3681, faculty of engineering sciences and techniques, IDIAP laboratory electrical and electronic engineering section, federal institute of technology in lausanne for obtaining the grade of doctor of sciences through EPF microtechnology engineer
17. Poulos M et al (2002) Neural network based person identification using EEG features. In: 2002 IEEE international conference on acoustics, speech, and signal processing. proceedings. ICASSP99 (cat. no.99ch36258). https://doi.org/10.1109/ICASSP.1999.759940
18. Alneam JS (2006) Brain tumors images diagnosis using hybrid intelligence techniques. A PH.D. thesis, submitted to the council of the college of Computers and Mathematics Science University of Mosul
19. Mandal S, Saha G, Pal RK (2015) Neural network training using firefly algorithm. Global J Adv Eng Sci (GJAES) 1(1) (ISSN: 2395–1001)
20. Yang XS (2009) Firefly algorithms for multimodal optimisation. In: Watanabe O, Zeugmann T (eds) Fifth symposium on stochastic algorithms, foundation and applications (SAGA 2009), LNCS, vol 5792, pp 169178

Traffic Sign Recognition Approach Using Artificial Neural Network and Chi-Squared Feature Selection

Manisha Vashisht and Brijesh Kumar

Abstract With the rising population and vehicular traffic across the globe, driver safety on road has become a huge concern for most governments. Emerging technologies and industrial revolutions have given rise to concept of autonomous cars. The driving systems embedded in these cars identify the traffic signs on the road and then take appropriate action. In spite of all these efforts, the accuracy of traffic sign image detection still remains a challenge for most car manufacturers and drivers, especially under difficult weather conditions. Multiple authors have done research in past and have proposed approaches relevant to identification of traffic sign images. The proposed solutions on traffic sign image detection have been influenced largely by Artificial Intelligence (AI)-based implementation techniques. In this research paper, authors have used Mapillary public traffic image dataset and have proposed an innovative approach using chi-squared ranking algorithm along with ANN for image classification. The effectiveness of proposed approach is compared with some related works. Experimental results showed that the proposed enhanced algorithm based on ANN and chi-squared algorithm provided better results.

Keywords Chi-squared test · Artificial neural network · Feature selection · Traffic sign image

1 Introduction

As per the research data from a safety survey, road accidents causing deaths and injuries across the globe, and human mistakes have been the root cause for more than 90% of the cases. Considering this data, it becomes much more relevant for research community to find ways using technological innovations to reduce road accidents caused by human errors [1, 2]. Research surveys have shown that, globally, road accidents have resulted in millions of people dying and suffering injuries every year. In October 2021, World Health Organization (WHO) has taken up a goal of

M. Vashisht (✉) · B. Kumar
Manav Rachna International Institute of Research and Studies, Faridabad, India
e-mail: manishavashisht@gmail.com

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_45

preventing 50% of the deaths and injuries caused by road accidents by the year 2030 [3]. In this paper, authors have made use of Mapillary image public dataset for conducting research on traffic sign images. The public dataset has repository of close to one lakh traffic sign images which have been gathered from across the globe, covering mostly possible weather conditions [4].

1.1 ANN

Historically, Artificial Neural Network has been known for its learning ability for pattern identification and has been widely used by authors in research related to solving problems in field of effort estimation, cost prediction, IoT-based traffic sign recognition, and software defects prediction [5–14]. In this research paper, authors have made use of ANN for image detection from public dataset of Mapillary Traffic Sign Images.

1.2 Feature Selection

Feature selection approach is used to reduce the unwanted and unnecessary dimensions of the large dataset. The objective is to focus on the limited number of meaningful dimensions to get a better prediction. In the proposed work, feature selection algorithm process is implemented for image recognition of traffic sign image dataset. Initially, ANN approach is utilized independently without implementing feature selection approach. Then in the next stage, both filter-based feature selection using chi-squared test and ANN technique are used as *n* hybrid approach for image detection.

Further, in the next section, authors have described the literature review in the related work. Proposed architecture and framework have been covered in Sect. 3, while Sect. 4 covers comparative analysis and overall results. The paper is concluded while describing the possible enhancements and expansion in the research area.

2 Literature Review

In this section, authors have described briefly about the past relevant research work that has been conducted on traffic sign image detection using ANN and chi-squared feature selection-based approaches and its results.

In this paper, authors have proposed feature selection-based approach for predicting the performance of students for devising strategies for educational initiatives. The execution results of the study showed that there is a variance of 10% in improved accuracy of prediction, when feature selection approach is used with only relevant dimensions [15].

German Traffic Sign Recognition Benchmarks (GTSRB) dataset was used by authors to train prediction model on a selective dataset of 43 classes using an integrated approach of implementing ANN and HOG (Histograms of Oriented Gradients, the approach yielded an accuracy of 80% [16].

In this paper, authors have proposed a novel approach using Internet of Vehicle-based artificial neural network (IoV-ANN) for monitoring and tracking the movement of vehicle. They used a global positioning system (GPS) for tracking and implemented ANN for classification. The IoV-ANN implementation results achieved the best performance of 97% and an error rate 9.12% [17].

Authors collected the urban accidents data of 2 years and made use of ANN to identify the most significant variables having an impact on the gravity of accidents and then arriving on the best possible of predicting accident. The execution results provided insights depicting multiple logistic regression with accuracy of 89% [18].

In this paper, authors have proposed an innovative approach for detecting and recognizing traffic signs. Initially, clustering algorithm has been used for pre-processing and then ANN is used during the cluster segmentation. ANN, Adaboost, and support vector machine (SVM)-based approach were validated with the newly introduced hybrid approach, and result showed that ANN-based approach outperforms the other two approaches [19].

In this paper, authors have suggested a pipeline-driven architecture for traffic sign image frame processing while collecting the video data acquired from camera by using AI. The system was tested under varying weather conditions for effectiveness of the proposed approach. Results showed precision of 89% during daytime and 36% at nights, indicating the improvement required for nighttime accuracy [20].

In this research work, authors have used an innovative approach using CNN to detect and recognize number plates of vehicles. The architecture consists of two parts; First part uses a camera for detecting number plate and extracting a higher resolution image for reconstructing the pixel image quality. In the other part, CNN is used to extract features from high resolution number plate image [21].

In this paper, authors have used chi-squared feature selection along with ANN classification to eliminate unwanted parameters from the landslide susceptibility dataset in order to optimize time required for processing the data and to enhance prediction accuracy. Results show chi-squared feature selection helps in bringing improvement in the overall accuracy caused by ANN model [22].

In the next section, authors have explained the ANN-based network model, architecture, and implementation of ANNs, feature section approach, and the investigational conclusions which assist in reaching an interpretation of traffic sign image detection accuracy.

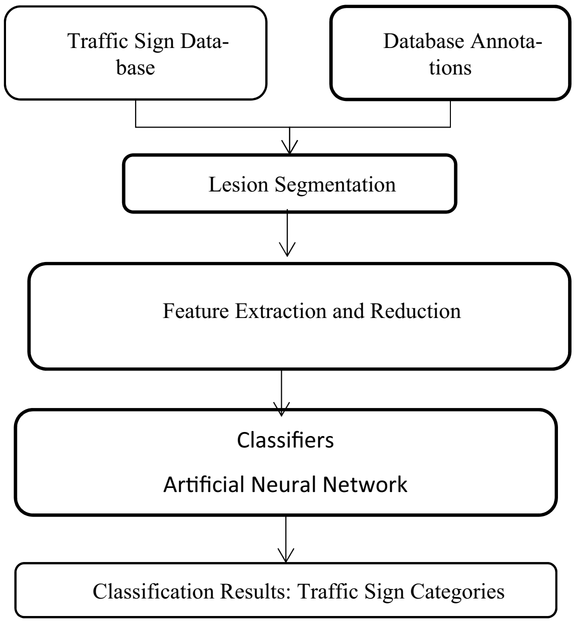
3 Proposed Framework for Model Design

The proposed framework has taken the images from the Mapillary Traffic Sign Image dataset. The dataset has close to around 5000 traffic sign images, which have been further categorized into 40 different categories, each category consisting of further 125 sign images. Every single traffic sign image further consists of 6528 specific data point dimensions. As a part of feature extraction, authors have used database annotations and HOG feature to find region of interest, which resulted in further reduction of feature parameters from 6528 data points to 372 data points.

In the first stage, implementation as per architecture diagram as shown in Fig. 1 is done, but without feature reduction technique and results are studied. Then, in the next stage, along with ANN, feature reduction using filter method-based chi-squared technique is used and the results are studied and compared. The proposed framework has been illustrated in Fig. 1.

Authors studied multiple past research work done in the similar field and finalized Levenberg–Marquardt (LM) method for neural network training. MATLAB scientific software was utilized, which provided neural network toolbox environment to code and execute the model. The software provided integrated development environment to configure the software to upload the image dataset and provided output result in the form of regression values. Figure 2 shows a configuration screen while execution of proposed model in the MATLAB environment.

Fig. 1 High level block diagram



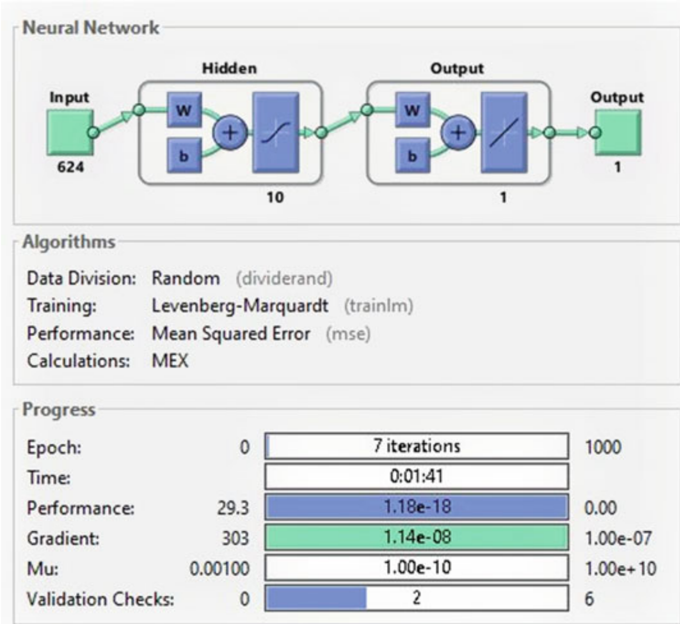


Fig. 2 ANN configuration

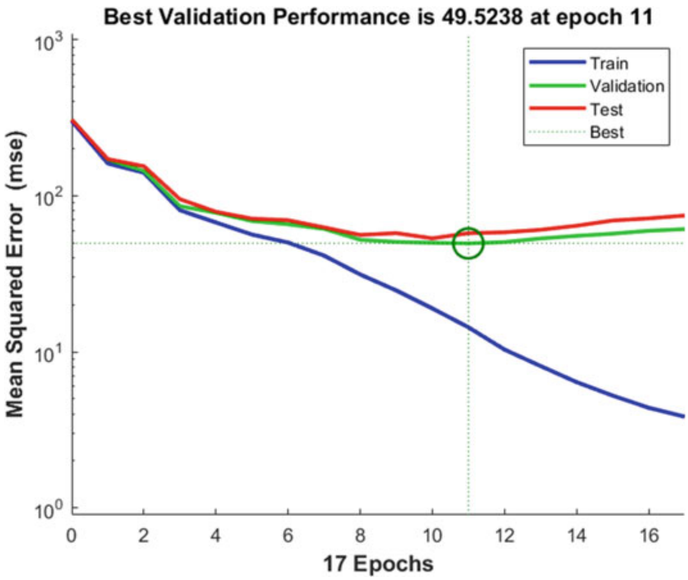


Fig. 3 Performance Plot of ANN without feature selection

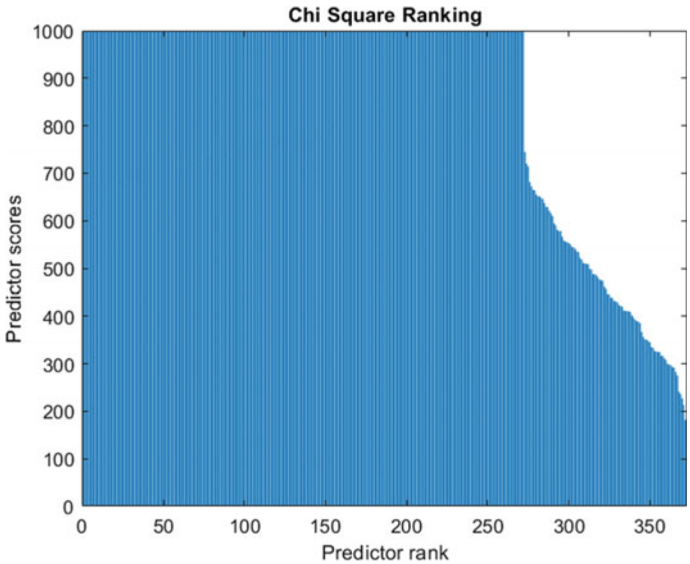


Fig. 4 Chi-squared ranking

Figure 3 represents the performance plot and depicts the best validation performance of 49.52 at epoch 11. `fscchi2` function in MATLAB validates the independence of input variable by utilizing explicit validation tests. Then, ranking of the input features takes place by making use of the p-values, which are made available as an output of the statistical tests performed as an execution on chi-squared algorithm.

As seen from Fig. 4, the predictor score is 1000 at 272 predictor ranks, which is further reduced to the level of 300 predictor rank, the feature points now selected are most relevant. Finally, 300 top features are selected from the overall 372 data points.

4 Results and Comparative Analysis

The obtained regression results describe a stronger relationship with regard to traffic sign image recognition. As evident from Figs. 5 and 6, the R value represents the strong correlation that has been existing between the output values and the target values. Test results where ANN along has been used as classifier with no ranking algorithm been used on the Mapillary traffic sign image dataset, an overall accuracy of 0.80 has been achieved by the proposed model. Further, the training dataset denotes R value of 0.84, while the validation dataset has achieved value of 0.75. Test results where ANN along has been used as classifier along with the chi-squared ranking algorithm been used on the Mapillary traffic sign image dataset, an overall accuracy of 0.89 has been achieved by the proposed model. Further, the training dataset denotes R value of 0.94, while the validation dataset has achieved value of 0.77. In ANN

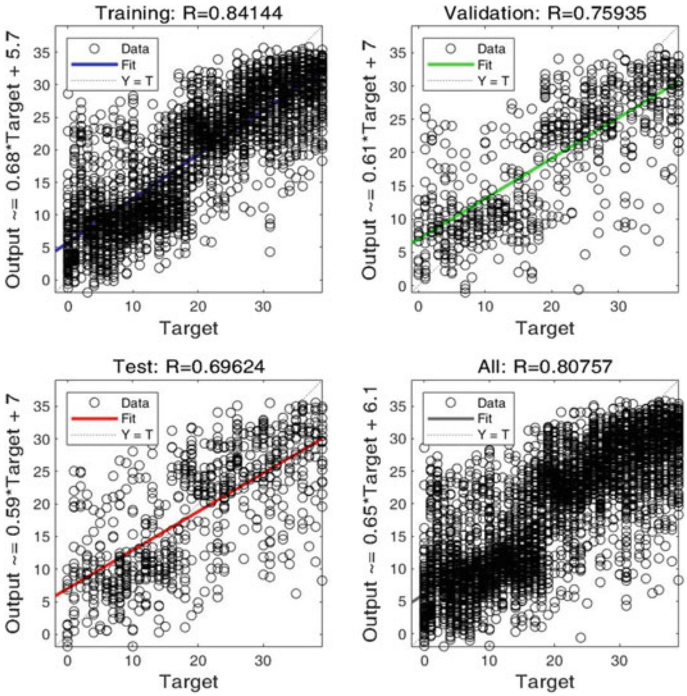


Fig. 5 Regression results without feature selection

classifier, there is an enhancement observed in the overall prediction accuracy results (Table 1).

5 Conclusion and Future Work Discussion

In this research paper, authors have proposed an innovative approach of utilizing feature selection approach using chi-squared approach integrated with ANN for traffic sign images detection from Mapillary public dataset. The results have been encouraging, and it has been established from the research that feature selection approach helps in reduction of redundant features and improves prediction accuracy. With respect to the scope of future research, it is suggested to validate the same approach on multiple image datasets and using other feature selection approaches.

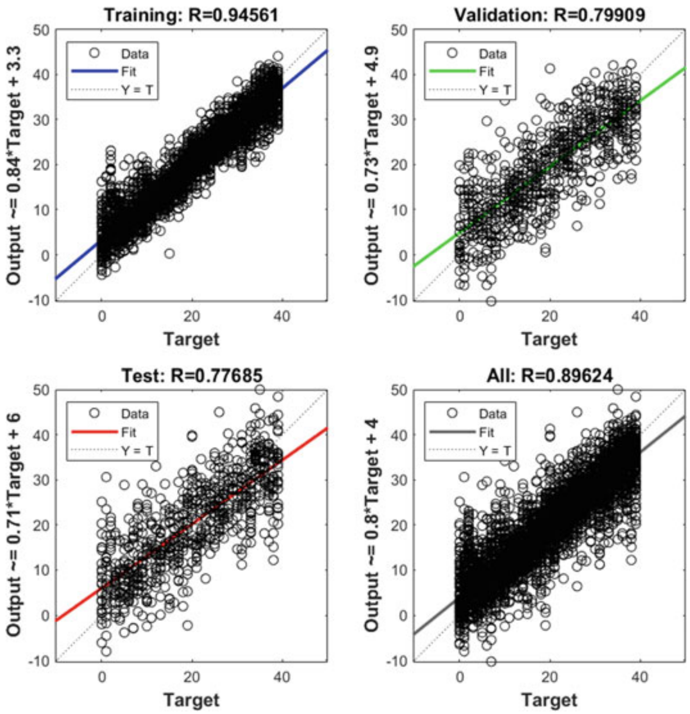


Fig. 6 Regression results with feature selection

Table 1 Comparative analysis

Reference model	Accuracy level (%)
Proposed model with ANN as classifier	80
Proposed model with ANN and chi-squared test for ranking	89

References

1. Traffic safety facts, research note. NHTSA's National Center for Statistics and Analysis, Apr 2019
2. Chand A, Jayesh S, Bhasi AB (2021) Road traffic accidents: an overview of data sources, analysis techniques and contributing factors. Mater Today Proc
3. Who.int. (2021) Second decade of action for road safety (online). Available at https://www.who.int/teams/social-determinants-of-health/safety-and-mobility/decade-of-action-for-road-safety-2021-2030. Accessed 28 Nov 2021
4. Ma D, Fan H, Li W, Ding X (2020) The state of Mapillary: an exploratory analysis. ISPRS Int J Geo Inf 9(1):10
5. Gopi R, Sathiyamoorthi V, Selvakumar S, Manikandan R, Chatterjee P, Jhanjhi NZ, Luhach AK (2021) Enhanced method of ANN based model for detection of DDoS attacks on multimedia internet of things. Multimedia Tools Appl:1–19

6. Nandal M, Mor N, Sood H (2021) An overview of use of artificial neural network in sustainable transport system. Comput Methods Data Eng:83–91
7. Usha V (2021) Traffic sign classification using deep learning. Turkish J Comput Math Educ (TURCOMAT) 12(9):250–253
8. Vashisht V, Lal M, Sureshchandar GS (2016) Defect prediction framework using neural networks for software enhancement projects. J Adv Math Comput Sci:1–12
9. Vashisht V, Lal M, Sureshchandar GS (2015) A framework for software defect prediction using neural networks. J Softw Eng Appl 8(08):384
10. Vashisht V, Kamya S, Vashisht M (2020) Defect prediction framework using neural networks for business intelligence technology-based projects. In: 2020 international conference on computer science, engineering and applications (ICCSEA). IEEE, pp 1–5
11. Vashisht V, Socis I (2016) An investigation into the various statistical process control tools. Int J Eng Technol 6(2)
12. Vashisht V. Measuring and analyzing the impact of implementing sub process monitoring and defect prediction model in the software development life cycle
13. Vashisht M, Kumar B (2021) Traffic sign recognition using multi-layer color texture and shape feature based on neural network classifier. In: Micro-electronics and telecommunication engineering. Springer, Singapore, pp 479–487
14. Vashisht M, Kumar B (2021) Effective implementation of machine learning algorithms using 3D colour texture feature for traffic sign detection for smart cities. Expert Syst:e12781
15. Zaffar M, Savita KS, Hashmani MA, Rizvi SSH (2018) A study of feature selection algorithms for predicting students' academic performance. Int J Adv Comput Sci Appl 9(5):541–549
16. Kerim A, Efe MÖ (2021) Recognition of traffic signs with artificial neural networks: a novel dataset and algorithm. In: 2021 international conference on artificial intelligence in information and communication (ICAIIC). IEEE, pp 171–176
17. Ghazal TM, Said RA, Taleb N (2021) Internet of vehicles and autonomous systems with AI for medical things. Soft Comput:1–13
18. Najafi Moghaddam Gilani V, Hosseinian SM, Ghasedi M, Nekoosa M (2021) Data-driven urban traffic accident analysis and prediction using logit and machine learning-based pattern recognition models. Math Probl Eng
19. Lahmyed R, El Ansari M, Kerkaou Z (2021) Automatic road sign detection and recognition based on neural network
20. Tchernykh A, Shepelev V (2021) Computational intelligence for analysis of traffic data. In: Smart cities: third Ibero-American congress, ICSC-Cities 2020, San José, Costa Rica, Nov 9–11, 2020, Revised selected papers, vol 1359. Springer Nature, p 167
21. Ahsan M, Based M, Haider J (2021) Intelligent system for vehicles number plate detection and recognition using convolutional neural networks. Technologies 9(1):9
22. Pham BT, Van Dao D, Acharya TD, Van Phong T, Costache R, Van Le H, Nguyen HBT, Prakash I (2021) Performance assessment of artificial neural network using chi-square and backward elimination feature selection methods for landslide susceptibility analysis. Environ Earth Sci 80(20):1–13

Formation Controller and Reinforcement Learning Algorithm in Multiple Surface Vessels

Dao Phuong Nam, Dang Van Trong, Pham Dinh Duong, and Nguyen Hong Quang

Abstract This brief presents a completed control structure, including formation control and adaptive reinforcement learning (ARL) algorithm for a multi-agent system of multiple surface vessels (SVs). The ARL strategy is established for each SV with the advantage of handling a non-autonomous system without solving the Hamilton–Jacobi–Bellman (HJB) equation. The additional formation controller is implemented to complete the control structure of multi-SV systems and guarantees the formation tracking problem. Simulation studies are developed to show the performance of the proposed control structure.

Keywords Surface vessel (SV) · Adaptive reinforcement learning (ARL) · Formation control · Multi-agent systems

1 Introduction

Surface vessel (SV) systems are the special case of robotic systems, and we can absolutely utilize the control structures in general robotics in considering the control design of SVs. Until recently, many control schemes have been proposed for robotic systems, for example, sliding mode control (SMC) technique for handling the trajectory control in Bilateral Teleoperators [1], a back-stepping technique in cascade controller of mobile robots [2–4], and Adaptive Reinforcement Learning (ARL)-based optimal control design in mobile robots [5–7]. It can be seen that the control methods can be generally categorized into two main groups, including classical nonlinear control techniques and optimal control approaches. However, almost all existing control structures for robotic systems have not been developed for multiple robotic systems [1–7], and they are rarely considered by optimal control. The optimality principle approach with the key idea to overcome these disadvantages

D. P. Nam · D. Van Trong · P. D. Duong
Hanoi University of Science and Technology, Hanoi, Vietnam

N. H. Quang (✉)
Thai Nguyen University of Technology, Thai Nguyen, Vietnam
e-mail: quang.nguyenhong@tnut.edu.vn

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_46

of solving Hamilton–Jacobi–Bellman and Ricatti equations is implementing the Adaptive Reinforcement Learning (ARL) technique [6]. On the contrary, multi-agent systems have been mentioned by many approaches [8–12]. In [10], the distributed control for multi-agent systems was presented with the consideration of the Kronecker product, Neural Networks, and Linear Matrix Inequalities (LMIs). Moreover, LQR optimal control was developed for multi-agent systems [9]. However, almost all previous control designs for multi-agent systems have considered each agent in easy cases of linear systems and sub-systems [8–12]. Additionally, optimal solutions for multi-agent systems have not mentioned the HJB equation's obstacle [8–12]. Unlike the previous control designs only focusing on each SVs or multiple SVs with traditional nonlinear control method, this paper proposed the control structure being the frame of formation control for generating the reference and ARL-based optimal control for each SV. The proposed algorithm can obtain tracking formation, trajectory tracking control, and the unification of optimality problem and stability.

2 Preliminaries and Problem Statement

In this section, we investigate a formation control of multiple Surface Vessels (SVs) with the following dynamic equation of each SV:

$$\begin{cases} \dot{q}_i = J(q_i)\upsilon_i \\ M(q_i)\dot{\upsilon}_i + C(\upsilon_i)\upsilon_i + g(q_i) = \tau_i + \Delta(q_i, \upsilon_i) \end{cases} \tag{1}$$

where $q_i = [x_i, y_i, \theta_i]^T \in \Re^3$ defines the vector of position, angle in the earth-fixed frame. $\upsilon_i = [u_i, \vartheta_i, r_i]^T \in \Re^3$ is a vector of a surge, sway, and yaw velocities in the body-fixed frame of each SV and $\tau_i \in \Re^n$ is the control input vector.

The control objective is to obtain the formation control with each SV controlled by ARL-based optimal control design. It is worth noting that the formation control designs the reference for each agent, and the advantage of this control design for each uncertain SV can implement an optimal controller based on the On-Policy iterative algorithm without solving the HJB equation.

3 Control System of Multiple Surface Vessels

This section focuses on establishing the completed control scheme with two parts, including Actor/Critic Online-based ARL algorithm for each SV and the formation controller generates appropriate reference for each agent.

First, Online Adaptive Reinforcement Learning for each Surface Vessel can be considered by utilizing the property of Hamiltonian and computation of δ_{hjb}. For each SV in Eq. 1, we proposed the following controller:

$$\tau_i = u_i + \tau_{id} \tag{2}$$

where

$$\tau_{id} = M(\upsilon_i)\frac{\mathrm{d}}{\mathrm{d}t}\upsilon_{id} + C(\upsilon_{id})\upsilon_{id} + D(\upsilon_{id})\upsilon_{id} + g(q_i) \tag{3}$$

with υ_{id} to be established in next sections and we implement ARL algorithm to find the term of u_i. The model in (1) can be transformed into the following time-invariant systems:

$$\frac{\mathrm{d}}{\mathrm{d}t}X_i = \begin{bmatrix} -M^{-1}l_1 + M^{-1}\left(\upsilon_{id}\left(z_{qi}, q_{id}\right)\right) \\ J\left(z_{qi} + q_{id}\right)z_{\upsilon i} - \beta_{qi}z_{qi} \\ h_1(q_{id}) \end{bmatrix} + \begin{bmatrix} M^{-1} \\ 0 \\ 0 \end{bmatrix}(u_i + \Delta_i) \tag{4}$$

The optimal purpose is to find u_i obtaining the minimum of performance index:

$$J(X_i, u_i) = \int_0^\infty \left(X_i^T Q X_i + u_i^T R u_i\right)\mathrm{d}\tau \tag{5}$$

where:

$$X_i = \left[z_{\upsilon i}^T, z_{qi}^T, q_{id}^T\right]^T \tag{6}$$

Similar to the work in [6], the online ARL-based optimal control is proposed as follows:

$$u_i^*(X_i) = -\frac{1}{2}R_i^{-1}G_i^T(X_i)\left(\frac{\partial\psi}{\partial x}\right)^T \hat{W}_{ai} \tag{7}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{W}_{ci} = -k_{ci}\lambda_i\frac{\sigma_i}{1 + \sigma_i^T\lambda\sigma_i}\delta_{hjb\_i} \tag{8}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{W}_{ai} = -\frac{k_{ai1}}{\sqrt{1 + \sigma_i^T\sigma_i}}\frac{\partial\psi}{\partial X_i}G_iR_i^{-1}G_i^T\frac{\partial\psi^T}{\partial X_i}\left(\hat{W}_{ai} - \hat{W}_{ci}\right)\delta_{hjb\_i} - k_{ai2}\left(\hat{W}_{ai} - \hat{W}_{ci}\right) \tag{9}$$

Second, for square formation control, this section presented the method to obtain an appropriate reference as follows:

$$
\begin{aligned}
\upsilon_i &= [\cos\theta_i, \sin\theta_i] \sum_{j\in N_i} \left(p_j - p_i - p_j^* + p_i^* - a^*\right) \\
&= [\cos\theta_i, \sin\theta_i](-(L\otimes I))\left(p_i - p_i^* - a^*\right) \\
\omega_i &= [-\sin\theta_i, \cos\theta_i] \sum_{j\in N_i} \left(p_j - p_i - p_j^* + p_i^* - a^*\right) \\
&= [-\sin\theta_i, \cos\theta_i](-(L\otimes I))\left(p_i - p_i^* - a^*\right)
\end{aligned} \tag{10}
$$

4 Simulation Results

The proposed algorithm is evaluated for a group of four SVs with the relation of them to be described by corresponding Laplacian matrix $L = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & -1 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & -1 & 2 \end{bmatrix}$ and the initial state of each SV are chosen as $q_1(0) = [0, 0, 0]^T$; $q_2(0) = [20, 0, 0]^T$; $q_3(0) = [10, 0, 0]^T$; $q_4(0) = [-10, 0, 0]^T$; $\upsilon_i = (0) = [0, 0, 0]^T$ $(\forall i = 1, 2, 3, 4)$.

On the contrary, the parameters of ARL algorithm for each SV are chosen as: $\beta_q = 2, k_c = 10, k_{a1} = 10, k_{a2} = 20, \nu = 0.01, Q = I_3, R = 1$.

According to the proposed control scheme in the previous section, the convergence of Weights in Actor/Critic RL algorithm (Fig. 1) and the tracking of formation (Figs. 2 and 3) describe the proposed method's effectiveness. The weight's convergence pointed out the complement of optimal control design of each SVs being obtained after learning process in finite time.

5 Conclusions

In this paper, the formation control for Multiple SVs combining with ARL-based optimal control for each agent is presented. This proposed method's advantage is to handle the time-varying systems in trajectory tracking control of each SVs by considering the corresponding time-invariant system. Moreover, the additional formation control scheme combining with ARL-based controller for each SV guarantees the tracking of desired formation. Finally, the simulation results verified the effectiveness of proposed algorithm in the situation that four uncertain SVs are controlled.

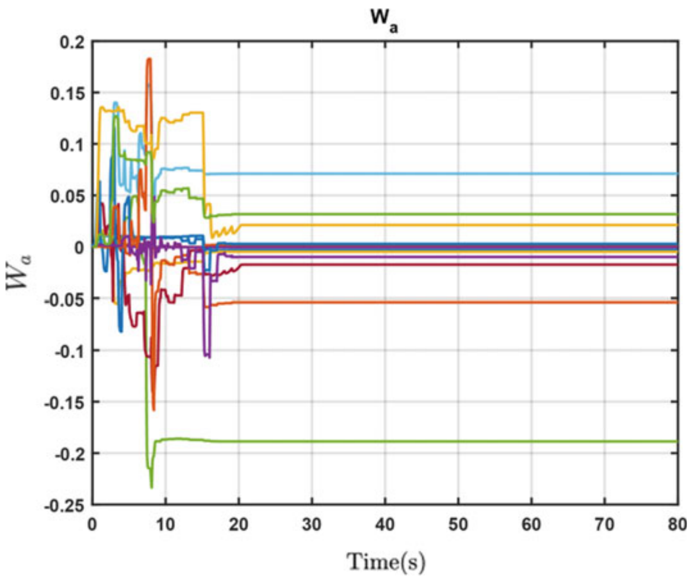
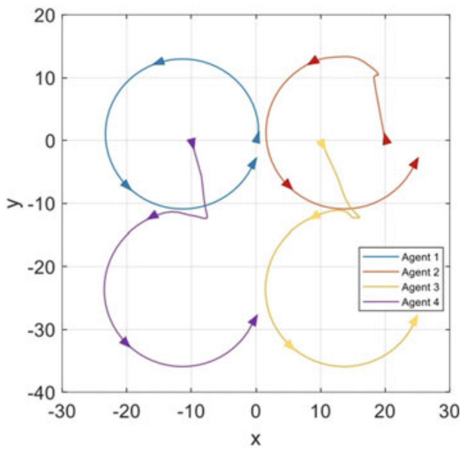


Fig. 1 The trained weights of actor part

Fig. 2 The formation tracking of multiple SVs (case 1)



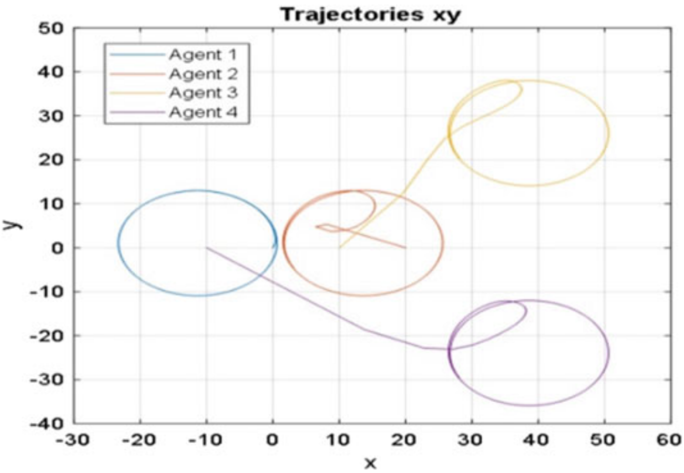


Fig. 3 The formation tracking of multiple SVs (case 2)

Acknowledgements This research was supported by the Thai Nguyen University of Technology.

References

1. Liu Y, Dao PN, Zhao KY (2020) On robust control of nonlinear teleoperators under dynamic uncertainties with variable time delays and without relative velocity. IEEE Trans Industr Inf 16(2):1272–1280. https://doi.org/10.1109/TII.2019.2950617
2. Dao PN, Nguyen HQ, Ngo MD, Ahn SJ (2020) On stability of perturbed nonlinear switched systems with adaptive reinforcement learning. Energies 13(19):5069
3. Binh NT, Tung NA, Nam DP, Quang NH (2019) An adaptive backstepping trajectory tracking control of a tractor trailer wheeled mobile robot. Int J Control Autom Syst 17(2):465–473
4. Nguyena T, Hoang T, Pham M, Dao N (2019) A Gaussian wavelet network-based robust adaptive tracking controller for a wheeled mobile robot with unknown wheel slips. Int J Control 92(11):2681–2692
5. Dao PN, Liu YC (2020) Adaptive reinforcement learning strategy with sliding mode control for unknown and disturbed wheeled inverted pendulum. Int J Control Autom Syst:1–12
6. Dao PN, Loc PT, Huy TQ (2021) Sliding variable-based online adaptive reinforcement learning of uncertain/disturbed nonlinear mechanical systems. J Control Autom Electr Syst:1–10
7. Dao PN, Nguyen HQ, Nguyen TL, Mai XS (2021) Finite horizon robust nonlinear model predictive control for wheeled mobile robots. Math Probl Eng
8. Sun D, Wang C, Shang W, Feng G (2009) A synchronization approach to trajectory tracking of multiple mobile robots while maintaining time-varying formations. IEEE Trans Rob 25(5):1074–1086
9. Movric KH, Lewis FL (2013) Cooperative optimal control for multi-agent systems on directed graph topologies. IEEE Trans Autom Control 59(3):769–774
10. Peng Z, Wang D, Zhang H, Sun G (2013) Distributed neural network control for adaptive synchronization of uncertain dynamical multi-agent systems. IEEE Trans Neural Netw Learn Syst 25(8):1508–1519

11. Zhao S, Dimarogonas DV, Sun Z, Bauso D (2017) A general approach to coordination control of mobile agents with motion constraints. IEEE Trans Autom Control 63(5):1509–1516
12. Hu X, Wei X, Gong Q, Gu J (2020) Adaptive synchronization of marine surface ships using disturbance rejection without leader velocity. ISA Trans

Data Dissemination in Vehicular Edge Network

Shilpi Mittal and Kanta Prasad Sharma

Abstract With the advancement of cloud computing and edge computing in traditional vehicular network, it has improved the QoS in VANET. When these heterogeneous networks use resources of clouds for computing and communication, it faces an issue of anxiety and latency due to long distance from the location of data generation. To concern this issue, we have implemented a data dissemination scheme compared with cloud and edge computing services in vehicular network. In this paper, architecture design with static and dynamic vehicle movement is illustrated and a data dissemination scheme is proposed on real-time scenario with the testing of QoS using simulator. The proposed scheme shows better result in terms of low latency and enhanced throughput.

Keywords Vehicular edge computing · VANET · Data dissemination · Vehicular networks · Cloud computing

1 Introduction

Nowadays, the emergent technologies in vehicular network are leading to improve the access of massive amount of data generated by smart vehicles after every second. To achieve this objective, various applications are developed and enhanced by Intelligent Transportation System (ITS). ITS is providing various framework for enhancing communication, computing, and storage in various types of vehicular network such as Vehicular Ad hoc Network (VANET) [1], Vehicular Cyber Physical System (VCPS) [2], Vehicular Cloud Computing (VCC) [3], Vehicular Fog Computing (VFC) [4], and Software Defined Vehicular Network (SDVN) [5]. All these techniques are providing an environment for efficient and reliable communication between vehicles in vehicular network as shown in Fig. 1.

VCPS is the combination of computation, inter-network communication, and physical processes in the vehicle for improving economic growth by reducing fuel

S. Mittal · K. Prasad Sharma (✉)
Chandigarh University, Mohali, Punjab, India
e-mail: tokpsharma@gmail.com

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_47

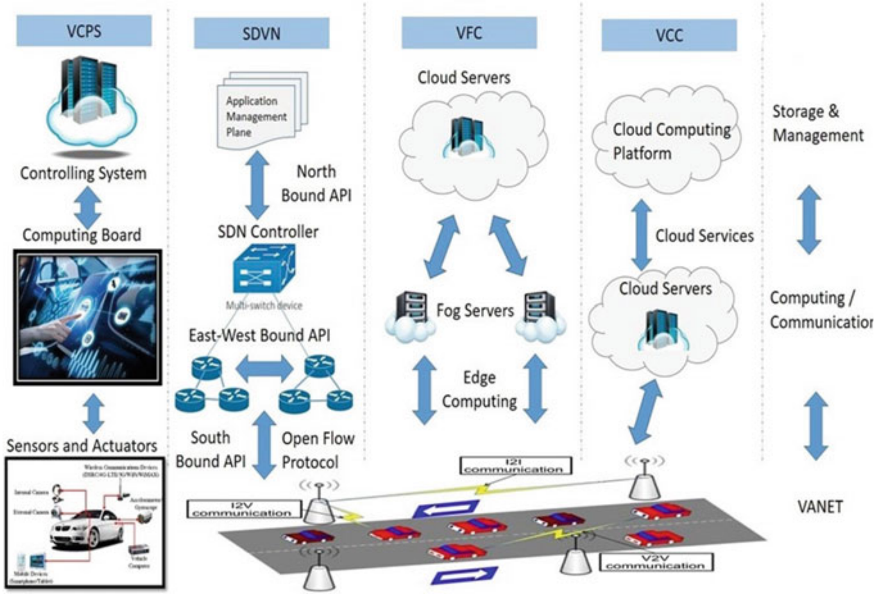


Fig. 1 VANET framework

consumption while driving. In VCC, vehicles on road communicate with each other through cloud servers. VCC and VCPS have improved the vehicular communication in terms of storage, communication, and computation but still it requires to improve latency and connectivity between nodes. To solve this issue fog computing has been integrated with conventional vehicular networks. In fog computing, fog servers compute data and deploy it to cloud servers by getting data from vehicle either via road side units (RSU) or on-board units (OBU) [6].

The advantage of fog computing is that it provides low latency, location awareness, and short distance local connection between vehicles which is a desirable characteristic for highly dynamic networks such as vehicular network. In SDVN, switches are deployed where data is generated and control communication to manage applications. Even though the advent of VANET has enhanced the applicability of transportation system, there are various research issues that have to be addressed for making the implementation of VANET a reality. There are various extensive studies which have gone to signify the development of different service architectures and communication protocols in traditional vehicular network. Vehicular networks are now extending themselves to revolutionary wireless technologies for improving communicating and computing in vehicular networks [5].

The major objectives of VANET are to maintain node stability when mobility, traffic safety, reliability, on-time end-to-end delivery and driving comfortably. An efficient data dissemination protocol helps in increasing the applicability of VANET, and so data dissemination protocols are one of the most important research challenges for these networks. As this paper has more concern on VANET, VCC, and VFC communication issues, some of the important data dissemination protocols based on these vehicular networks have been reviewed in this section.

This paper is concern to provide reliable communication by including the appropriate edge servers between the vehicles and cloud servers. These edge servers can be computing vehicles or RSU servers for intermediate communication. In the proposed work, a framework is introduced that included edge servers as intermediate commuters. This proposed framework will enhance communication at local vehicular network nearby user vehicle. Extensive simulation is conducted using simulator. The main contribution of the proposed work is as follows:

- We propose vehicular edge architecture that will facilitate communication in moving and static vehicles.
- A noval data dissemination scheme is proposed, which provides efficient and reliable communication services to overcome the latency.
- Proposed data dissemination scheme is implemented on real vehicular network scenario with mobility and without mobility.
- Finally, we conduct extensive simulations and results that the proposed scheme compared with existing schemes.

The organization of the paper is as follows: Literature survey is mentioned in Sect. 2. Section 3 includes the system model for proposed time varying traffic prediction scheme. Section 4 defines the simulation environment and result analysis. Section 5 concludes the work done.

2 Literature Review

Some of the literature based on VANET illustrate the vehicular communication-based solutions. Dua et al. [7] proposed ReIDD protocol to solve the problem of network congestion caused by message flooding in the network. To achieve higher data delivery, Tian et al. [8] has proposed an TrAD protocol which is applicable for both urban and highway scenarios. Ahmad et al.[9] proposed CODIE scheme which is a controlled data packet propagation to control the broadcast storm in vehicular network. Bi et al. [10] proposed UMBP with an objective of emergency message dissemination in urban transportation system. Wu et al. [11] proposed ADDSEN middleware which is a cyber physical sensing framework for managing distributed knowledge in drone swarm urban areas.

Ucar et al. [12] proposed a hybrid architecture VMaSC-LTE with the objective of this framework to provide higher efficiency in vehicular Networks. To handle high mobility, a multi-hop clustering algorithm VMaSC has been proposed based on relative mobility metric and overhead of cluster connection. He et al. [13] proposed ODDA algorithm based on the expansion of dimensions and the concept of dynamic programming. The key objective is to deploy dropboxes with optimum deployment cost in between a single pair of source and destination in dense vehicular networks.

Kim et al. [14] have proposed cloud system-based approach. This technique follows deterministic approach for data retrieval from connected vehicles in the networks. The route-based data prefetching method has been designed.

Khawatreh et al. [15] has proposed IHVCDT model to satisfy the data access through wireless channels having different data types. Paranjothi et al. [16] have proposed DFCV approach by emerging two technologies, one is fog computing, and another is Cloud computing. Aazam et al. [17] has proposed MeFoRE methodology to provide resource estimation and improve Quality of Service (QoS) at fog computing with VANET. Xu et al. [18] has proposed a DRAM to balance resource allocation in fog environment to avoid overload and bottleneck problems while computing Internet of Things (IoT) applications at fog edges. Huang et al. [19] has formalized vehicular fog computing for vehicular networks as the extension of fog computing paradigm.

3 System Model

Vehicles are directly connected to each other through V2V, V2I, or V2R communication. Therefore, as vehicles are equipped with enormous sensors they generate lots of data and also receive huge amount of data vice versa. In the proposed work, it is assumed that vehicles are equipped with smart sensors and act as smart vehicle. Vehicles are static at a parking lot and forming a vehicular network. It has been also assumed the parking area is covered in the range of radio frequency with the use of WiMax towers. To extend this network, cloud is included in the proposed network and also deployed edge between cloud and vehicles. Edge is initiated by cloud, and vehicles are controlled by cloud by sharing vehicle ID, location, and stability in network for time T.

In the proposed network, vehicles send data to edge server and edge server send back required result to source vehicle. As explained in Algorithm 1:

Algorithm 1 proposed Vehicular edge process

Step 1: Begin
Step 2: Initialize a vehicular network including edge and cloud
Step 3: Vehicles are sending data through broadcast
Step 4: Edge node receive data and forward to cloud for future use
Step 5: Interested vehicle receive data and send acknowledgment
Step 6: Extend information to next coming vehicles and reset the process
Step 7: End

The communication between vehicles under VEC framework has shown in Fig. 2. Vehicular communication is divided into two different scenarios vehicle-to-vehicle (V2V) edge network and vehicle-to-infrastructure (V2I) edge network. In V2V edge network, vehicle itself acts as client or edge node. Vehicles disseminate information through smart OBU. One vehicle acts as client and offloads task to nearest vehicle

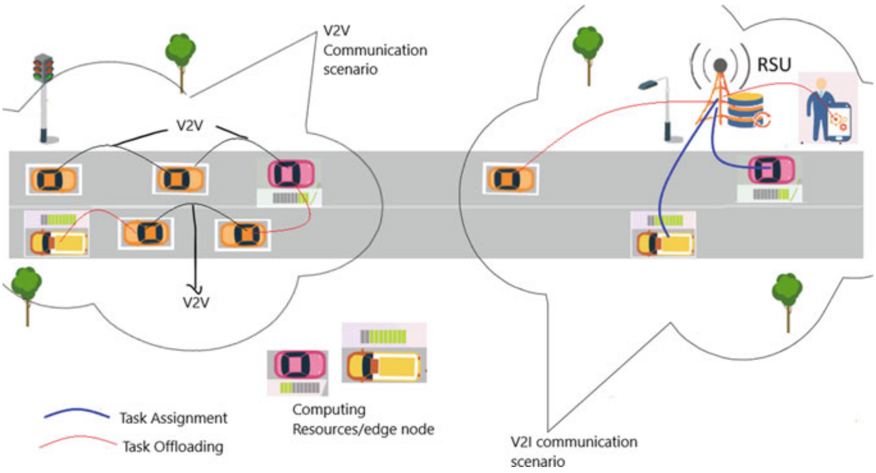


Fig. 2 Vehicular edge network communication model

for computation. That computing vehicle acts as server which stores and computes data for further dissemination, whereas in V2I scenario, a server edge is installed beside RSU to get request for task computation.

Once data is offloaded to edge, it may compute task at its own or distribute among vehicles. Therefore, vehicles can extend the communication to next joining edge vehicles. The objective of vehicular edge computing is to achieve communication, computing, and resource availability nearest to the user vehicle. Thus, it is the key role of VEC to address the requirements of exponentially increasing data generated by vehicles and the use of edge devices on low-delay and high bandwidth.

Now forwarding to the working principle of proposed model. The list of variables is given in Table 1, and the process is described below.

Table 1 Variables of algorithms

Variables	Description
$TTL(Vn)$	Vehicles's time-to-live
$Exp(TTL(Vn))$	Expected time-to-compute
vn	Velocity of vehicle
an	Acceleration of vehicles
PEg	Primary edge server node
SEg	Secondary edge server node
RCm	Remote cloud server
Cn	Edge user's task computation time
δm	Computation time of remote cloud server
N	Total number of vehicles
K	Total number of edge servers

To initiate task computing, very first step is how to initialize edge network for task distribution and computation. It is assumed that N autonomous vehicle flow on road and centralized cloud server have the information about flow of vehicles. When these moving vehicles stand on stopping places for fixed time T, where T denotes the time for how long vehicle will wait to cross the road on intersection point or to go ahead on its way. At this time, client vehicle sends request to edge vehicle PEg for its initialization. Then, edge vehicle sends confirmation message for instance activation to the client vehicle. Now, edge vehicle will create edge network by sending a joining message to other vehicles. As the vehicles accept joining request, they become part of vehicular edge network. Now, edge node assigns task to these edge vehicles in distribution manner. When computation task is over, edge nodes revert back computed result to PEg edge node, and finally, PEg node collects these responses and forwards result to requested client vehicle.

Algorithm 2 Edge initialization Algorithm

Begin
$i \leftarrow 0$, $j \leftarrow 0$
for each Vmatrix M_i **do**
    $PE \leftarrow M_i$; //Transmit task to Primary-Edge
    $Edge_j \leftarrow PE$; //Primary edge become source
    for each Edge $Edge_j$ **do**
        $Edge_j \leftarrow Edge_{j+1}$; // Edge nodes coordinates and start computation
        Edge nodes switch to each-other
    end for //Edge executes the task
    $J = 0$;
    Next node offload task for computation
end for
Stop //Computation process completed

During the process shown in Algorithm 2, it might be possible that edge vehicles may leave edge network without completing its task; then, it is important to keep the edge network constant so that computation load can handle constantly. To maintain edge network, there are some considerations to follow to retain edge network consistent. If suppose there were K edge nodes in the network and M edge nodes leave the network after T time , then the size of edge network is below in (1) and the remaining nodes in the network is calculated as below:

$$K = K - M, N1 = N - K - M \tag{1}$$

else check the distance of vehicle V from PEg and add into edge network.

4 Simulation Results and Evaluation

To enhance the optimization of communication among heterogeneous vehicular network, simulation has been done on NS3 and SUMO. A real static scenario listed in Table 2 has been designed and simulated for communication testing at a fixed vehicular location. Vehicles are idle and can be used as edge nodes; thus, edge is included, and comparative analysis is done between edge, cloud, and VANET communication.

As expected, edge computing topology is showing better results in terms of throughput, less round trip time, and delta time because of the dedicated servers for each wireless network. Most of the stations in the cloud computing topology experienced high latency and less throughput because of limited channel bandwidth and total number of stations. There are few parameters are evaluated such as latency, throughput and packet loss. These parameters are compared and shown in table and graphs.

Latency is very less comparative to cloud and VANET. It means as edge is closest to vehicle and is powerful to store and compute comparative to cloud, it has improved the result with minimum latency. In terms of transmission rate, in VEC it is very high as comparative to cloud and VANET.

Simulation result depicts that VEC is better in transmission of packets with minimal latency as compared to other schemes.

Table 3 describes the comparative result in packet transmission among edge computing, cloud computing, and VANET. VANET has generated maximum packets and has highest connection-oriented transmission. There is equal number of packets lost and re-transmitted in edge computing. Cloud computing and VANET have identical maximum throughput, whereas edge computing has highest minimum throughput. There is very less latency in edge computing which is the greatest advantage of it (Table 4).

Figure 3 illustrates the flow of data transmission in comparative three schemes. VANET has identical transmission rate, whereas in cloud computing environment it

Table 2 Simulation parameters

Simulation parameters	
Parameters	Values
Phy mode	OFDM
Data rate	6 mbps
Packet size	160 bytes
No. of packets	20,100
Interval	1 s
No. of vehicles	8, 16, 24, 32, 64
No. of edge nodes	4, 8, 12, 16, 20
Communication protocol	V2V, V2I

Table 3 Comparison of different attributes

	Edge computing	Cloud computing	VANET
Total packets	3,293,675	4,200,985	4,294,923
TCP Retransmission (packets)	294	5973	8586
Lost packets	294	1794	2822
Maximum throughput (Mbps)	10.8	10.7	10.7
Minimum throughput (Mbps)	9.57	2.63	1.06
Maximum latency	1.03	3.0	4.01

Table 4 Result analysis

Edge parameters	Values
Lost packets	294
Maximum throughput (Mbps)	10.8
Minimum throughput (Mbps)	9.57
Maximum latency	1.03

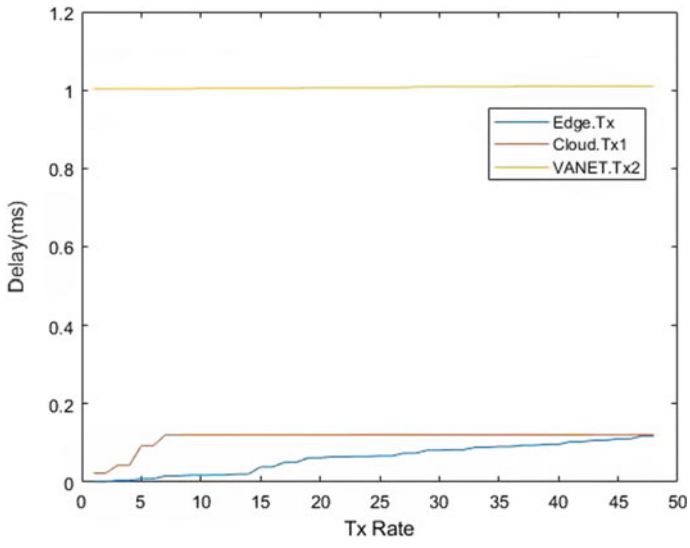


Fig. 3 Data transmission in VEC, VCC, VANET

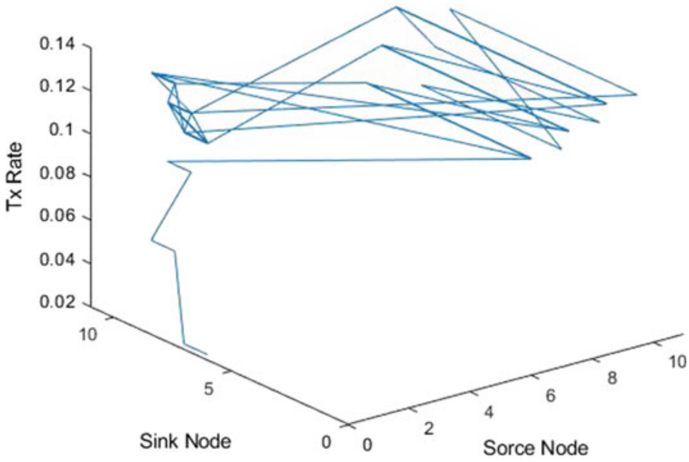


Fig. 4 Communication in VCC

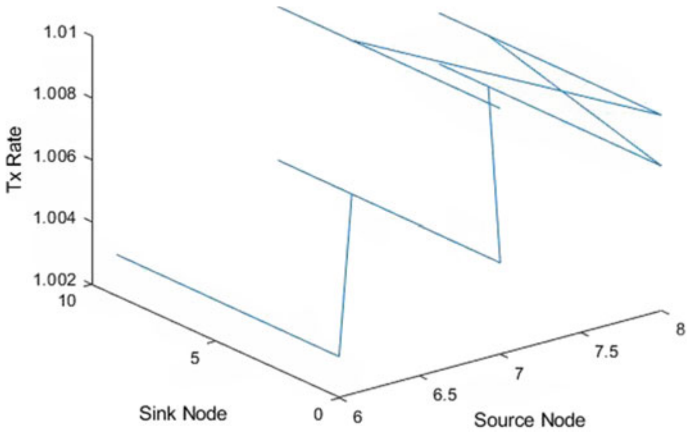


Fig. 5 Communication in VANET

increased at starting point but then become stable. Transmission rate in edge computing is slightly increasing throughout the communication and steady at 0.1 maximum.

Fig. 6 shows the transmission rate in VEC communication model. It shows that when sink nodes are increasing continuously, the transmission rate is getting high steadily, and in Fig. 4, transmission rate is increased when source nodes are getting increased. In VANET data transmission is constant as shown in Fig. 5.

The above mentioned simulation results show that the proposed data dissemination scheme in vehicular network is efficient and proved lower latency and high throughput. The implementation of edge computing in framed real-time scenario proved that communication through edge is better than direct communication through cloud computing; the comparative result shows that there is efficient communication

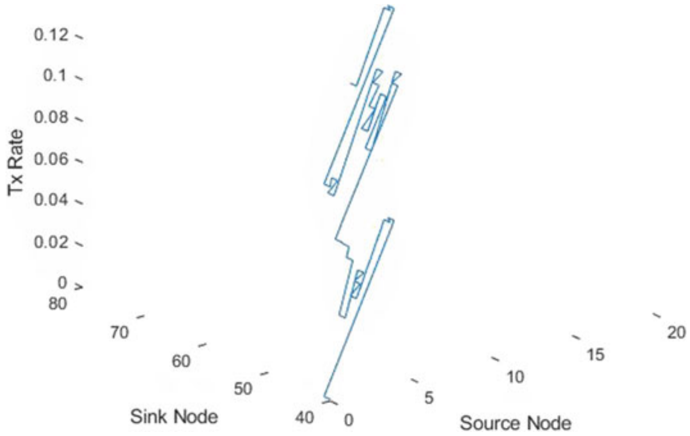


Fig. 6 Communication in VEC

between vehicular nodes when they disseminate information through nearest edge vehicle node. Hence, the proposed architecture enhanced the reliable communication between vehicles in vehicular network.

5 Conclusion

The proposed scheme is describing the VANET with including edge for communication. It also presented a communication model between edge, cloud, and VANET application deployment. VANET communication along with edge support instead of cloud is showing better communication result through the simulation in NS3. The work can be extended by including more testing on communication with hybrid vehicular network technologies. The work will be extended in future for improving vehicular communication in heterogeneous vehicular network for better throughput and lesser latency.

References

1. Hartenstein H, Laberteaux K (2009) VANET: vehicular applications and inter-networking technologies, vol 1. Wiley, Hoboken
2. Rajkumar R, Lee I, Sha L, Stankovic J (2010) Cyber-physical systems: the next computing revolution. Design automation conference. IEEE, New York, pp 731–736
3. Caveney D (2010) Cooperative vehicular safety applications. IEEE Control Syst Maga 30(4):38–53

4. Lytrivis P, Thomaidis G, Tsogas M, Amditis A (2011) An advanced cooperative path prediction algorithm for safety applications in vehicular networks. IEEE Trans Intell Transp Syst 12(3):669–679
5. Lee E, Lee EK, Gerla M, Oh SY (2014) Vehicular cloud networking: architecture and design principles. IEEE Commun Maga 52(2):148–155. https://doi.org/10.1109/MCOM.2014.6736756
6. Toor Y, Muhlethaler P, Laouiti A, De La Fortelle A (2008) Vehicle ad hoc networks: applications and related technical issues. IEEE Commun Surveys Tutorials 10(3):74–88
7. Dua A, Kumar N, Bawa S (2017) Reidd: reliability-aware intelligent data dissemination protocol for broadcast storm problem in vehicular ad hoc networks. Telecommun Syst 64(3):439–458
8. Tian B, Hou KM, Zhou H (2016) The traffic adaptive data dissemination (TrAD) protocol for both urban and highway scenarios. Sensors 16(6):920
9. Ahmed SH, Bouk SH, Yaqub MA, Kim D, Song H, Lloret J (2016) Codie: controlled data and interest evaluation in vehicular named data networks. IEEE Trans Veh Technol 65(6):3954–3963
10. Bi Y, Shan H, Shen XS, Wang N, Zhao H (2015) A multi-hop broadcast protocol for emergency message dissemination in urban vehicular ad hoc networks. IEEE Trans Intell Transp Syst 17(3):736–750
11. Wu D, Arkhipov DI, Kim M, Talcott CL, Regan AC, McCann JA, Venkatasubramanian N (2016) ADDSEN: adaptive data processing and dissemination for drone swarms in urban sensing. IEEE Trans Comput 66(2):183–198
12. Ucar S, Ergen SC, Ozkasap O (2015) Multihop-cluster-based IEEE 802.11p and LTE hybrid architecture for vanet safety message dissemination. IEEE Trans Veh Technol **65**(4), 2621–2636
13. He J, Ni Y, Cai L, Pan J, Chen C (2017) Optimal dropbox deployment algorithm for data dissemination in vehicular networks. IEEE Trans Mob Comput 17(3):632–645
14. Kim R, Lim H, Krishnamachari B (2015) Prefetching-based data dissemination in vehicular cloud systems. IEEE Trans Veh Technol 65(1):292–306
15. Khawatreh SA, Al-Zubi EN (2017) Improved hybrid model in vehicular clouds based on data types (ihvcdt). Int J Adv Comput Sci Appl 8(8):114–118
16. Paranjothi A, Khan MS, Atiquzzaman M (2018) DFCV: a novel approach for message dissemination in connected vehicles using dynamic fog. International conference on wired/wireless internet communication. Springer, Berlin, pp 311–322
17. Aazam M, St-Hilaire M, Lung CH, Lambadaris I (2016) MeFoRE: QoE based resource estimation at Fog to enhance QoS in IoT. In: 2016 23rd International conference on telecommunications, ICT 2016. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICT.2016.7500362
18. Xu X, Xue Y, Li X, Qi L, Wan S (2019) A computation offloading method for edge computing with vehicle-to-everything. IEEE Access 7:131068–131077. https://doi.org/10.1109/ACCESS.2019.2940295
19. Huang C, Lu R, Choo KKR (2017) Vehicular fog computing: architecture, use case, and security and forensic challenges. IEEE Commun Maga 55(11):105–111

A Deep Learning Architecture for Human Activity Recognition Using PPG and Inertial Sensor Dataset

Rohit Kumar Bondugula, Kaushik Bhargav Sivangi, and Siba K. Udgata

Abstract Human activity recognition helps identify the activity of a person based on data provided by sensors. The wireless wearable sensors provide robust techniques for data collection and classification. Most wearable devices contain heart rate and body orientation detection sensors. In this work, we experimented on a Photoplethysmography sensor used for heart rate identification and accelerometer signals to recognize the subject's orientation to classify the human activities. We proposed a novel deep learning architecture (MiniVGG) which was able to find the right activity time interval that resulted in overall lower false positives and false negatives. Our proposed model, MiniVGG, gave the highest accuracy of **97.75%** on the PPG dataset higher than any other existing models. The results of our experiments are compared with other baseline models and at varied sampling time window sizes and have shown greater accuracy. In addition, we report the best combination of the sampling time window size and the appropriate model to achieve the best accuracy, minimum false positives, or minimum false negatives depending on the requirement. This helps in developing a multi-criteria decision-making system for human activity recognition system using wearable sensor devices.

Keywords Human activity recognition · Deep learning · MiniVGG · Photoplethysmography

1 Introduction

In computer vision, human activity recognition (HAR) is one of the most active research areas for various contexts like security, skill assessment [1], and industrial settings [2], smart homes [3, 4], surveillance [5], healthcare [6, 7], and human–computer Interaction. Human activity can be tracked using wearable sensors, and smart devices directly positioned on the human body where it will help in tracking the activity of a particular person doing.

R. K. Bondugula · K. B. Sivangi · S. K. Udgata (✉)
School of Computer and Information Sciences, University of Hyderabad, Hyderabad, India
e-mail: udgata@uohyd.ac.in

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_48

Photoplethysmography (PPG) is a physiological practical technique to find human activity with help of other parameters approach in wearable devices. To measure the change in the volume of blood in the microvascular tissue, the employed device will release and make reflect light from the device and observes the different rates. The photo sensor reads the reflected light read by photo sensor to detect the changes. The valid heart rate estimation can be obtained from the output of the sensors whilst the PPG is intraoperative, non-invasive for HR estimation corresponding to the electrocardiography (ECG) and surface electromyography.

Nonetheless, accurate evaluation of the PPG signals can be taken down from the wrist when the subject is performing various activities often challenging problems. Due to the relative moment between PPG light detector and wrist skin, PPG signal might be corrupt the accurate information. Various signal processing techniques are proposed on data derived by sensor types and accelerometer [8, 9] data to reduce the motion artefacts will be useful.

Smartphones and smartwatches are becoming the most popular and data fusion techniques of PPG can be used for providing accurate reliable information. Amidst smartphones, smart electronic wearables built-in triaxle accelerometers are probably the most popular sensors [10]. PPG sensors are not designed to capture the motion signals as opposed to IMU, typically g accelerometers and gyroscopes so not applied in HAR classification. But, the PPG sensor is very useful for HAR [11] like because PPG-enabled smartwatches and wristbands are becoming more common and almost always include a PPG sensor, it makes sense to take advantage of the information it can provide, especially since it comes at no additional cost to the user of one of these PPG-enabled devices.

The PPG sensor can be used alone when other HAR sensors are unavailable or combined with them to improve recognition performance. This sensor may be used to track a variety of physiological characteristics (heart rate, blood volume, and so on) all in one place. As a result, we decided to use the PPG signal to predict human actions as well. HAR consists of machine learning techniques that have proven that the HAR is treated as a pattern recognition problem. When it is about deep learning, the HAR can be categorized as conventional machine learning techniques and deep learning-based techniques. Various machine learning methods like K-nearest neighbours, support vector machines, Gaussian mixture models, hidden Markov models [12], random forests [13], and molecular complex detection [14] methods were adopted.

In recent things of machine learning algorithms and portable device hardware cover for replaceable for wearables and allows implementing deep learning algorithms directly on embedded microcontrollers without the need for transferring data to system with limited computational power and low-energy consumption. In recently, edge computing appeared to reduce communication latency, network traffic, communication cost, and privacy concerns, and edge devices cannot support high-computation loads. As, we discussed a lot of machine learning DNN [15] models have been developed for HAR. Particularly, the deep learning algorithms proved with high performance with high computation, which makes them insufficient to deploy on edge devices. The rest of paper is discussed as follows: In Sect. 1.1, we have discussed the literature review in detail. The dataset description is given in Sect. 2. In

Sect. 3, the proposed methodology is explained in detail. In Sect. 4, a comprehensive discussion on the experimental results and analysis is explained, which is followed by the performance comparison and discussions in Sect. 4.3. Finally, in Sect. 5, we discussed the conclusions and future work.

1.1 Literature Review

Deep learning brought great improvements in signal recognition and object detection. Zhang et al. [16] worked on the recognition in the infrared images based on CNN along with automatic target detection. Using player identity, a deep learning identification model for players identification, 2D localization, and segmentation based on a Cascade mask R-CNN model implemented [17] where player identity id is ignored in existing methods with help of robust multi-camera multi-player tracking framework.

Most of the researchers explicated the deep learning techniques and adopted the classification of PPG. Variation between consecutive heartbeats measured in milliseconds known as heart rate variability will helpful for cardiovascular diseases like acute myocardial infarction and irregular heartbeat. The biLSTM is the betterment of long short-term memory (LSTM), and main motto is to collect the information behind and ahead of specific sample points by receiving the forward and backward loads. To derive the three indexes for HRV estimation of accurate PPG cardiac period segmentation, RNN-based biLSTM was introduced in [18].

To estimate blood pressure ECG and PPG signals combined with RNN structure. The biLSTM is used as the input hidden layer to look for contextual features both forward and backward in [19]. Different CNN architectures for PPG-based heart rate estimation are studied in [20].

Bangaru et al. [21] investigated that many researchers developed the kinematic sensor-related worker activity recognition techniques with the considerable prediction outcome. He identified key challenge from the state-of-art techniques which utilizes smartphones to the practical implementation, less classified activities, and minimal recognized motions and their body parts. To overcome this challenge, he proposed an artificial neural network which mainly discussed about data fusion and data acquisition. The architecture was validated on the scaffdd data which gives accuracy rate of 94%.

Elshafei et al. [22] proposed a feed-forward neural network to identify the impact of biceps muscle fatigue in the human activity recognition. The proposed architecture validated on the real-time dataset which contains 3000 biceps concentration curls performed and collected from the volunteers whose age is in between 20 and 35. He finally concluded that changes were often in data patterns during the fatigue presence which cause features to statistically insignificant.

Brophy et al. [23] worked on the wrist-worn-based smart devices which can able to access the insights which are present in the human behaviour and as well as health based on the sophisticated analytics. Bommi et al. [24] proposed a novel WCM model for the early classification. The objective of his work to increase the accuracy of prediction by implementing a weighted consensus model.

Photoplethysmography is a common method to detect the human activity recognition in the wearable devices [25]. Most of the PPG signals are corrupted by motion-based artefacts. Alessandrini [25] addressed the human activity by introducing the recurrent neural network in low-cost, low-power micro-controller by maintaining the accuracy rate of prediction more efficiently. The recurrent neural networks give more accuracy as compared to most of the state-of-art techniques. The implemented architecture is effective for the wearable devices to predict human activity recognition.

Aydemir et al. [26] proposed a robust technique for the PPG segments by Hibert transform and then classified by Naive Bayes, nearest neighbours, and decision tree algorithms. He concluded that human activity information can able to monitor heart rates and early screenings of different atherosclerotic pathologies.

2 Dataset Description

The dataset [8] included with this paper is useful for analyzing the PPG signal obtained from the wrist using the Maxim Integrated MAXREFDES100 sensor as shown in Fig. 1. The dataset comprises of recorded data acquired from 7 subjects which includes 105 PPG signals (15 for each participant) and the associated 105 tri-axial accelerometer signals measured at a sampling frequency 400 Hz. For training and testing, a publicly available dataset [8] was used for our experiment. Seven adults volunteered to participate in data collection exercises. The data recorded with along with the activity can be seen in Table 1.

A specific weight lifting cuff as seen in Fig. 1 was used to apply the sensor device effectively to the wrist. It is flexible by a tear-off closure and has excellent elastic properties, making it especially suitable for ensuring a perfect adherence of the sensing device to the surface of the skin. The sensor was then initially fastened on the subject's wrist, with the cable (used in the "tethered" mode) coming out from the

Table 1 Data consistency: each subject's data acquisition time

Subjects	Squat activity [s]	Stepper activity [s]	Resting activity [s]
1	311.5975	442.9900	3271.7
2	216.7975	397.6150	2962.8
3	231.4950	271.0400	1323.8
4	212.5750	269.6800	1361.9
5	246.2950	241.9750	1440.9
6	237.3700	325.9025	1402.0
7	266.8600	254.9300	1510.7

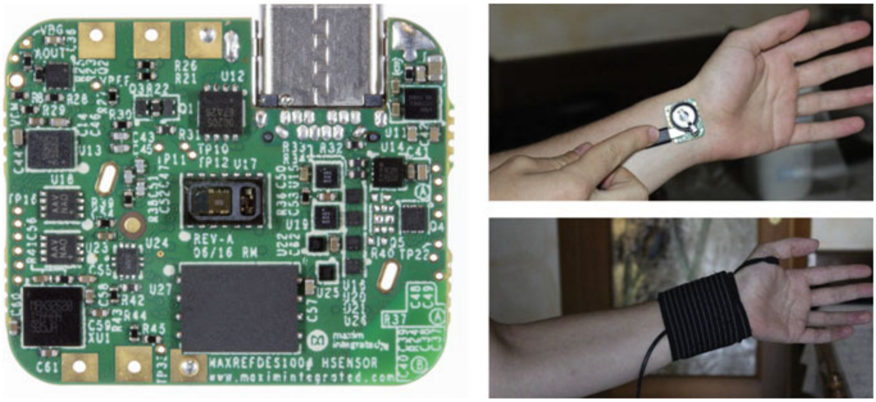


Fig. 1 Maxrefdes100 PPG sensor and its placement for data acquisition

back end of the band, which was then suitably tightened. The activities in the dataset are divided into 3 classes as follows:

1. Squats.
2. Stepper.
3. Resting.

3 Proposed Methods

3.1 MiniVGG

The VGGNet was initially developed for image-based datasets and was first applied on ImageNet. It is based on convolutional neural networks stacked one upon each other to increase the depth of extracting the features. Based on this idea of stacking multiple convolutional layers, we developed MiniVGG for time series data. This architecture is a simple deep network consisting of multiple convolutional layers along with intermediate pooling layers that help in reducing the dimensionality of the features. Further, the output from these convolutional layers is passed to fully connected layers to predict the activity. The complete architecture of MiniVGG is depicted in Fig. 2.

The input is initially given to a stack of two 1-dimensional convolutional layers. The number of feature maps within each layer is set to 224 and 192, respectively, whereas, the kernel size is set to 5, 3, respectively, for each layer. The output from this stack is passed on to a 1-dimensional max pooling layer. The pool size is set to default as 2, and the same is followed for all the max pooling layers. The pooling layer reduces the dimensionality of the features from convolutional layer to half. These new features are passed to another stack of convolutional layers followed by

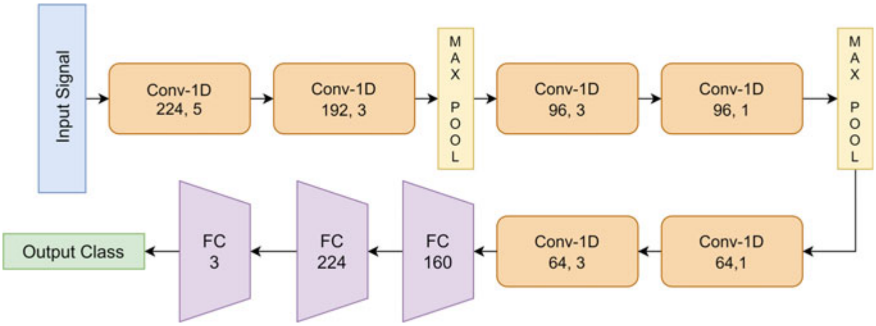


Fig. 2 MiniVGG architecture

a max pooling layer. In this stack, both the convolutional layers generate 96 features maps. The kernel size is set to 3, 1, respectively. The low-dimensional representation is passed to two more convolutional layers both generating 64 feature maps. This depth in the network allows for extraction of multiple unseen patterns underlying in the input signal in the initial layers and identify them as the network deepens. After the complete feature extraction from the convolutional layers, the features are passed through multiple fully connected layers. Specifically, we used two fully connected layers consisting of 160 and 224 hidden units. ReLU activation has been used for both the layers. Finally, the output layer consists of 3 units (number of output classes).

For comparing the performance of the proposed method, MiniVGG, we also implemented the original VGG-16 method for sequential data. Keeping the hyperparameters intact as the original model, we changed the 2-dimensional convolutional layers to 1-dimensional as meant for the input sensor signal. The architecture of the VGG-16 model is depicted in Fig. 3. A total of 13 convolutional layers are utilized to build the the model. The initial 2 layers consist of 64 feature maps each with a kernel size set to 3. As the network progresses, the feature maps are doubled at every stage as represented. A max pooling layer is used between stacks of convolutional layers to reduce the dimensionality. Finally, two fully connected layers with 4096 units are used with ReLU activation function, and the final output layer consists of 3 units with softmax activation for output classification.

4 Experimental Results and Analysis

4.1 Results of Proposed MiniVGG Model

The dataset consists of PPG and tri-axial accelerometer signals. The signals are measured against three activities, namely rest, squat, and stepper exercises. The complete dataset is modified to split the signals based on window samples to generate multiple samples. For our experimentation, we used window sizes ranging from

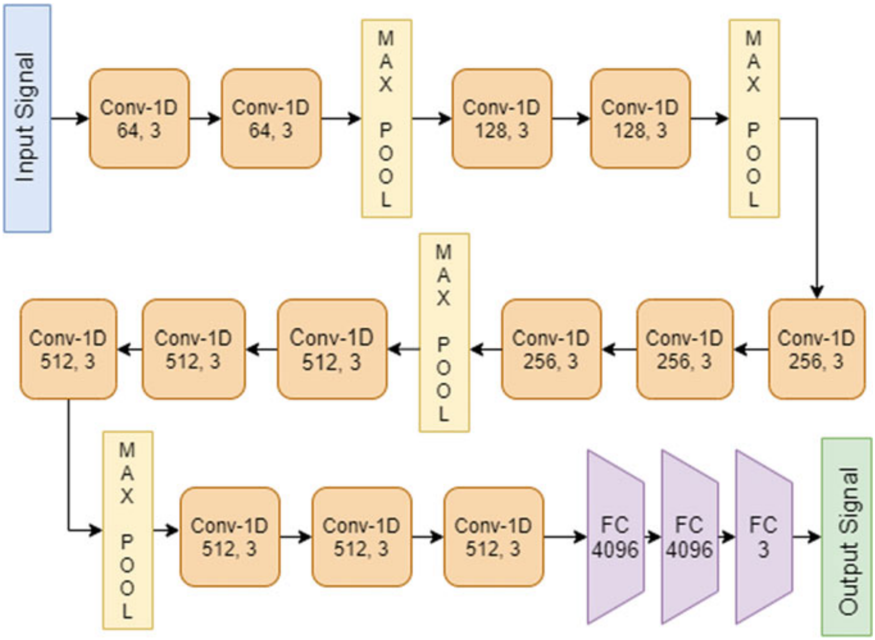


Fig. 3 VGG-16 architecture

1 to 5 s. This is because various activities require various time intervals for the heart beat to drastically increase/decrease, therefore we applied variable sizes to detect such irregularities. We developed the MiniVGG model consisting of multiple convolutional layers which can detect multiple patterns of the underlying signals and use the features to identify the activity being performed. The performance of the MiniVGG model on the human activity data is represented in Table 2.

The MiniVGG model has shown test accuracies greater than **97%** over all the window sizes. It has shown the highest accuracy of **98.92%** over 4 s window interval. As shown in Table 5, the MiniVGG model has shown lower false positives and false negatives over 4 s window interval. It was clearly able to distinguish squat activity from rest of the classes in this interval. This is evident from the fact that at lower window sizes, a higher number of false positives and false negatives were reported between squat and stepping activity. This is because, a small time interval is insufficient to identify the irregularities produced from PPG signal. The confusion matrices corresponding to all the intervals are depicted in Fig. 4. The classes belonging to step and squat activity show larger false positives and false negatives compared to the rest activity. This is due to the fact that less data are available belonging to these classes compared to the rest activity. Compared to standard VGG-16 model, MiniVGG has shown better results overall at all lower time intervals.

Table 2 Results of the MiniVGG model with corresponding window size and activities

Model	W.S	Activity	Precision	Recall	F1	Specificity	Sup	Train acc	Test acc.
MiniVGG	1	Rest	0.99	0.99	0.99	0.99	456	97.88	97.75
		Squat	0.97	0.96	0.97	0.99	285		
		Step	0.91	0.94	0.93	0.95	368		
		Avg	**0.95**	**0.96**	**0.96**	**0.98**	**1109**		
MiniVGG	2	Rest	0.99	0.99	0.99	0.99	425	98.05	96.85
		Squat	0.98	0.91	0.95	0.99	274		
		Step	0.93	0.98	0.96	0.96	380		
		Avg	**0.97**	**0.96**	**0.97**	**0.98**	**1079**		
MiniVGG	3	Rest	0.99	0.99	0.99	0.99	452	99.14	98.57
		Squat	0.99	0.96	0.98	0.99	241		
		Step	0.97	0.99	0.98	0.99	356		
		Avg	**0.98**	**0.98**	**0.99**	**0.99**	**1049**		
MiniVGG	4	Rest	0.98	1.00	0.99	0.98	441	99.28	98.92
		Squat	0.99	1.00	0.99	0.99	246		
		Step	1.00	0.97	0.98	1.00	332		
		Avg	**0.99**	**0.99**	**0.99**	**0.99**	**1019**		
MiniVGG	4	Rest	1.00	0.99	0.99	1.00	457	99.17	98.79
		Squat	0.97	0.98	0.98	0.99	200		
		Step	0.98	0.98	0.98	0.99	332		
		Avg	**0.98**	**0.98**	**0.98**	**0.99**	**989**		
Overall Avg			**0.97**	**0.97**	**0.98**	**0.99**	**5245**	**98.70**	**98.17**

Table 3 Results of the VGG model with corresponding window sizes and activities

Model	W.S	Activity	Precision	Recall	F1	Specificity	Sup	Train acc.	Test acc.
VGG	1s	Rest	1.00	0.99	0.99	1.00	456	95.74	95.04
		Squat	0.92	0.88	0.90	0.97	285		
		Step	0.91	0.94	0.93	0.95	368		
		Avg	**0.95**	**0.94**	**0.94**	**0.97**	**1109**		
VGG	2s	Rest	0.99	1.00	0.99	0.99	425	95.36	92.96
		Squat	0.86	0.86	0.86	0.95	274		
		Step	0.90	0.89	0.89	0.94	380		
		Avg	**0.92**	**0.92**	**0.91**	**0.96**	**1079**		
VGG	3s	Rest	0.99	1.00	0.99	0.99	452	98.52	98.28
		Squat	0.97	0.95	0.96	0.99	241		
		Step	0.96	0.98	0.97	0.98	356		
		Avg	**0.97**	**0.98**	**0.97**	**0.99**	**1149**		
VGG	4s	Rest	1.00	1.00	1.00	1.00	441	99.16	98.82
		Squat	0.97	0.98	0.98	0.99	246		
		Step	0.98	0.98	0.98	0.99	332		
		Avg	**0.98**	**0.98**	**0.98**	**0.99**	**1019**		
VGG	5s	Rest	1.00	1.00	1.00	1.00	457	99.39	99.09
		Squat	0.99	0.96	0.98	0.99	200		
		Step	0.98	0.99	0.99	0.98	332		
		Avg	**0.99**	**0.98**	**0.99**	**0.99**	**989**		
Overall Avg			**0.96**	**0.96**	**0.96**	**0.98**	**5245**	**97.63**	**96.84**

4.2 Results of the VGG Model

The VGG model has shown an average test accuracy of **96.84%** as shown in Table 3. This is lower compared to the average test accuracy of the proposed MiniVGG model which showed about **98.17%**. When compared to the MiniVGG model's average accuracy, the VGG-16 has shown 1.36% lower average accuracy. However, as shown in Fig. 5, on the 5 s interval, VGG-16 has outperformed MiniVGG by showing a 0.3% improvement. This resulted in lower false positives and false negatives at that interval as shown in Table 4. Therefore, we can infer, for activities requiring lower time intervals to be identified, the MiniVGG model can be employed, whereas for activities requiring higher time intervals to accurately classify, the VGG-16 can slightly show better results at additional computation cost.

Table 4 False positives and false negatives of respective classes for each classifier on the PPG - 3 axes accelerometer

Window size	Model	Activity	FN	FP	Test samples	Test accuracy	Test time (s)
1s	MiniVGG	Rest	1	5	1109	97.75	0.291
		Squat	9	10			
		Step	15	10			
2s	MiniVGG	Rest	4	2	1079	96.85	0.382
		Squat	4	24			
		Step	26	8			
3s	MiniVGG	Rest	5	1	1049	98.57	0.467
		Squat	1	10			
		Step	9	4			
4s	MiniVGG	Rest	8	0	1019	**98.92**	0.538
		Squat	**3**	**0**			
		Step	0	11			
5s	MiniVGG	Rest	0	3	989	98.79	0.895
		Squat	6	3			
		Step	6	6			
1s	VGG	Rest	0	2	1109	95.04	0.318
		Squat	21	32			
		Step	34	21			
2s	VGG	Rest	1	0	1079	92.96	0.477
		Squat	38	37			
		Step	37	39			
3s	VGG	Rest	1	0	1049	98.28	0.665
		Squat	6	11			
		Step	11	7			
4s	VGG	Rest	0	0	1019	98.82	0.749
		Squat	7	5			
		Step	5	7			
5s	VGG	**Rest**	**0**	**0**	989	**99.09**	0.889
		Squat	1	8			
		Step	8	1			

4.3 Performance Comparison and Discussions

Table 5 reports the comparison of proposed methodologies against baseline models for human activity recognition. For the fair comparison on baseline models, we reported the results of our proposed methodologies over 1 s window interval. The proposed MiniVGG has shown the highest accuracy of **97.75%** which is a significant improvement over the best baseline model PBP by 1.36%. The standard VGG-16

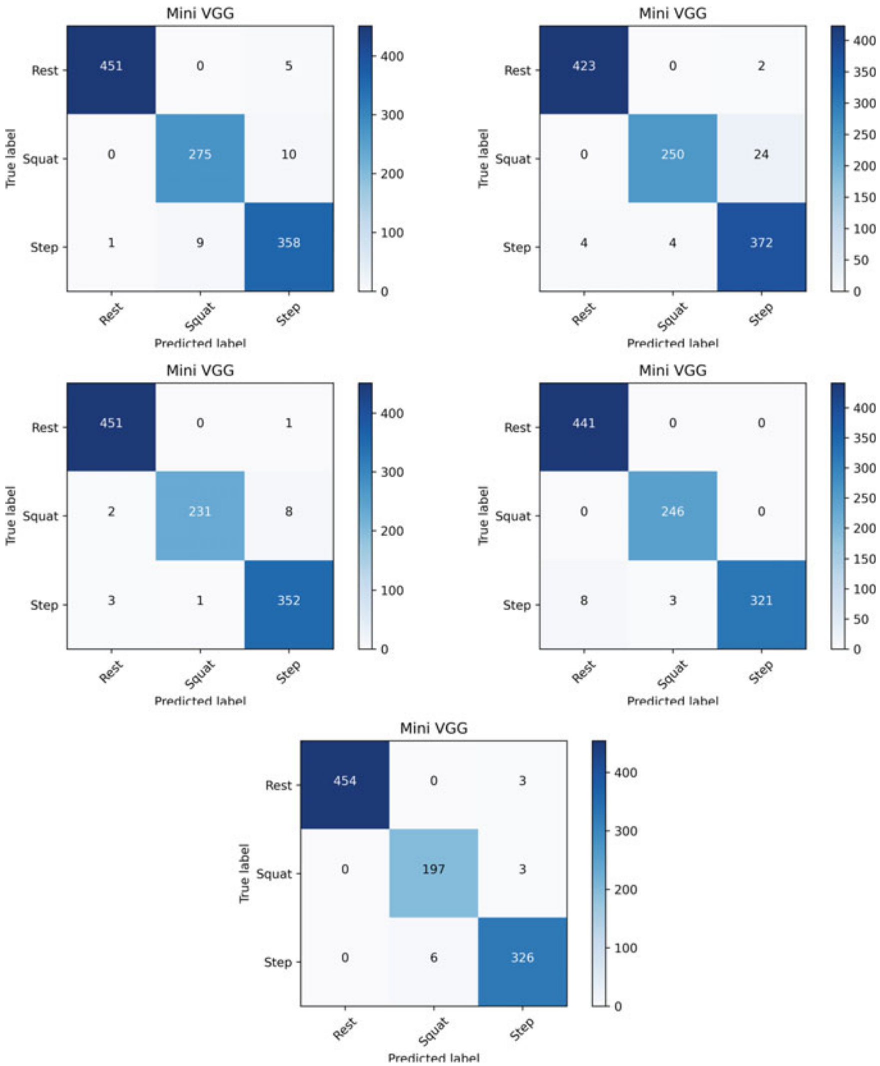


Fig. 4 Confusion matrices of the MiniVGG method on windowed samples

used as a baseline has shown lower accuracies compared to PBP and RNN baseline model. Based on Table 4, the standard VGG has the highest accuracy of 99.09% at 5 s window interval and has shown lower false positives and negatives.

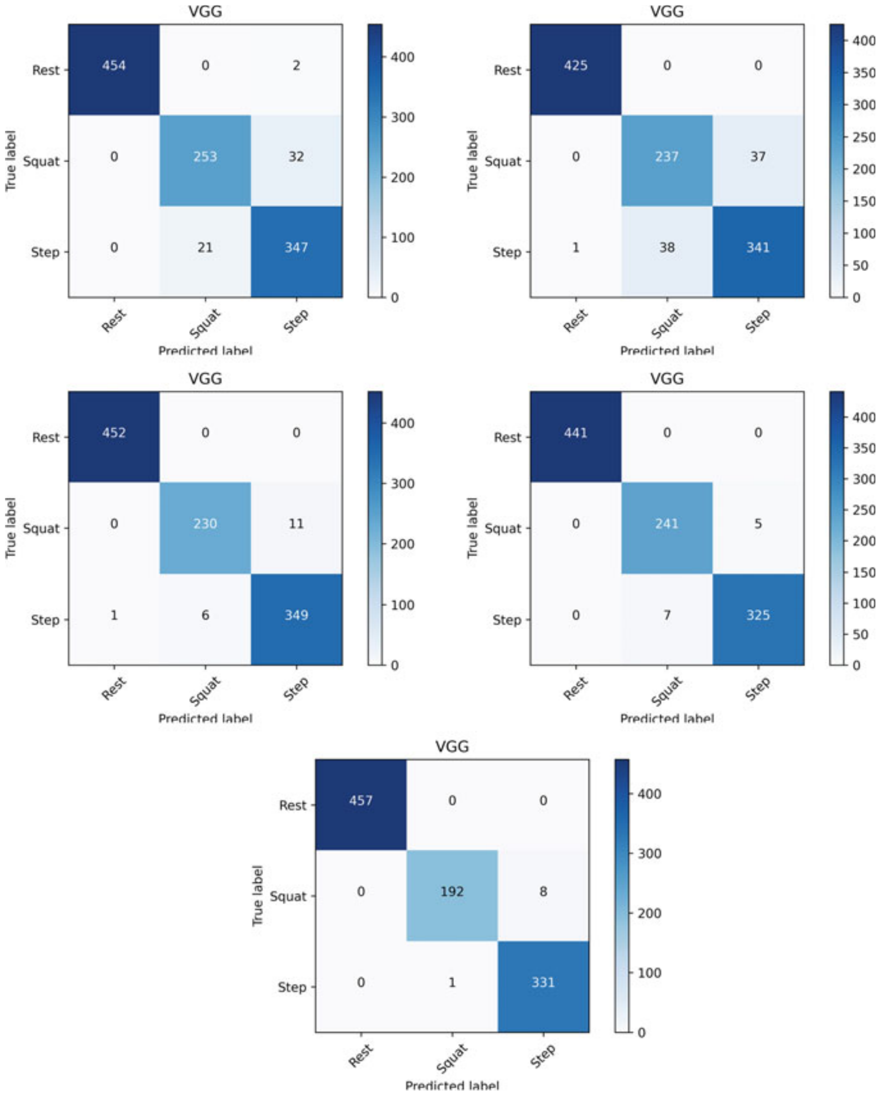


Fig. 5 Confusion matrices of the VGG-16 method on windowed samples

5 Conclusions and Future Scope

In this work, we propose a novel deep learning architecture, MiniVGG, based on VGG-16 model for human activity recognition. We employed the model on publicly available dataset consisting of PPG and tri-axial accelerometer signals. We categorized the data based on window sizes ranging from 1 s to 5 s to identify how the model

Table 5 Comparison of proposed methods with the existing methods

Model	Accuracy (%)
KLT + GMM [27]	78.00
PBP [28]	96.42
RNN/ LSTM [29]	95.54
MiniVGG	**97.75**
VGG-16	95.04

performs on varied time intervals for classifying activities. Our model has shown the best result of 98.92% and overall greater than 97% on all the time intervals. The proposed method has outperformed other baseline models applied on the same dataset and also has outperformed the proposed baseline VGG-16 model applied on sequential data. As shown in Tables 2, MiniVGG has shown an average recall of **97%** which shows the performance of the model to correctly identify the activity and an average specificity of **99%** indicating lower false positives. However, deeper networks suffer from the vanishing gradient problem. As described, the MiniVGG model consists of 3 blocks of two convolutional layers and further connected to 3 fully connected layer. There is a possibility of vanishing gradient occurring in this case, therefore in our future work, we would like to introduce skip connections between the blocks which might result in better accuracies and lower false positives and negatives.

References

1. Kranz M, Möller A, Hammerla N, Diewald S, Plötz T, Olivier P, Roalter L (2013) The mobile fitness coach: towards individualized skill assessment using personalized mobile devices. Pervasive Mob Comput 9(2):203–215
2. Stiefmeier T, Roggen D, Ogris G, Lukowicz P, Tröster G (2008) Wearable activity tracking in car manufacturing. IEEE Pervasive Comput 7(2):42–50
3. Cicirelli F, Fortino G, Giordano A, Guerrieri A, Spezzano G, Vinci A (2016) On the design of smart homes: a framework for activity recognition in home environment. J Med Syst 40(9):1–17
4. Rashidi P, Cook DJ (2009) Keeping the resident in the loop: adapting the smart home to the user. IEEE Trans Syst Man Cybernet-part A: Syst Hum 39(5):949–959
5. Chen L, Wei H, Ferryman J (2013) A survey of human motion analysis using depth imagery. Pattern Recogn Lett 34(15):1995–2006
6. Patel S, Park H, Bonato P, Chan L, Rodgers M (2012) A review of wearable sensors and systems with application in rehabilitation. J Neuroeng Rehab 9(1):1–17
7. Avci A, Bosch S, Marin-Perianu M, Marin-Perianu R, Havinga P (2010) Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: a survey. In: 23th International conference on architecture of computing systems 2010, 1–10, VDE
8. Biagetti G, Crippa P, Falaschetti L, Saraceni L, Tiranti A, Turchetti C (2020) Dataset from PPG wireless sensor for activity monitoring. Data Brief 29:105044
9. Zhang Z, Pi Z, Liu B (2014) TROIKA: a general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise. IEEE Trans Biomed Eng 62(2):522–531

10. Dernbach S, Das B, Krishnan NC, Thomas BL, Cook DJ (2012) Simple and complex activity recognition through smart phones. In: 2012 Eighth international conference on intelligent environments. IEEE, New York, pp 214–221
11. Boukhechba M, Cai L, Wu C, Barnes LE (2019) ActiPPG: using deep neural networks for activity recognition from wrist-worn photoplethysmography (PPG) sensors. Smart Health 14:100082
12. Attal F, Mohammed S, Dedabrishvili M, Chamroukhi F, Oukhellou L, Amirat Y (2015) Physical human activity recognition using wearable sensors. Sensors 15(12):31314–31338
13. Casale P, Pujol O, Radeva P (2011) Human activity recognition from accelerometer data using a wearable device. In: Iberian conference on pattern recognition and image analysis. Springer, Berlin, pp 289–296
14. Lu Y, Wei Y, Liu L, Zhong J, Sun L, Liu Y (2017) Towards unsupervised physical activity recognition using smartphone accelerometers. Multimedia Tools Appl 76(8):10701–10719
15. Walse KH, Dharaskar RV, Thakare VM (2016) PCA based optimal ANN classifiers for human activity recognition using mobile sensors data. In: Proceedings of first international conference on information and communication technology for intelligent systems, vol 1. Springer, Berlin, pp 429–436
16. Zhang R, Xu L, Yu Z, Shi Y, Mu C, Xu M (2021) Deep-IRTarget: an automatic target detector in infrared imagery using dual-domain feature extraction and allocation. IEEE Trans Multimedia
17. Zhang R, Wu L, Yang Y, Wu W, Chen Y, Xu M (2020) Multi-camera multi-player tracking with deep player identification in sports video. Pattern Recogn 102:107260
18. Xu K, Jiang X, Ren H, Liu X, Chen W (2019) Deep recurrent neural network for extracting pulse rate variability from photoplethysmography during strenuous physical exercise. In: 2019 IEEE biomedical circuits and systems conference (BioCAS). IEEE, New York, pp 1–4
19. Şentürk Ü, Yücedağ I, Polat K (2018) Repetitive neural network (RNN) based blood pressure estimation using PPG and ECG signals. In: 2018 2nd international symposium on multidisciplinary studies and innovative technologies (ISMSIT). IEEE, New York, pp 1–4
20. Reiss A, Indlekofer I, Schmidt P, Van Laerhoven K (2019) Deep PPG: large-scale heart rate estimation with convolutional neural networks. Sensors 19(14):3079
21. Bangaru SS, Wang C, Busam SA, Aghazadeh F (2021) ANN-based automated scaffold builder activity recognition through wearable EMG and IMU sensors. Autom Constr 126:103653
22. Elshafei M, Costa DE, Shihab E (2021) On the impact of biceps muscle fatigue in human activity recognition. Sensors 21(4):1070
23. Brophy E, Muehlhausen W, Smeaton AF, Ward TE (2020) Optimised convolutional neural networks for heart rate estimation and human activity recognition in wrist worn sensing applications. arXiv preprint arXiv:2004.00505
24. Bondugula RK, Udgata SK, Bommi NS (2021) A novel weighted consensus machine learning model for covid-19 infection classification using CT scan images. Arab J Sci Eng, pp 1–12
25. Alessandrini M, Biagetti G, Crippa P, Falaschetti L, Turchetti C (2021) Recurrent neural network for human activity recognition in embedded systems using PPG and accelerometer data. Electronics 10(14):1715
26. Aydemir T, Şahin M, Aydemir O (2020) A new method for activity monitoring using photoplethysmography signals recorded by wireless sensor. J Med Biol Eng 40(6):934–942
27. Moghadam ZB, Noghondar MS, Goshvarpour A (2021) Novel delayed Poincare's plot indices of photoplethysmogram for classification of physical activities. Appl Med Inf 43(1):43–55
28. Mahmud T, Akash SS, Fattah SA, Zhu W-P, Ahmad MO (2020) Human activity recognition from multi-modal wearable sensor data using deep multi-stage LSTM architecture based on temporal feature aggregation. In: 2020 IEEE 63rd International Midwest symposium on circuits and systems (MWSCAS). IEEE, New York, pp 249–252 (2020)
29. Brophy E, Veiga JJD, Wang Z, Ward TE (2018) A machine vision approach to human activity recognition using photoplethysmograph sensor data. In: 2018 29th Irish signals and systems conference (ISSC). IEEE, New York, pp 1–6

Reversible Watermarking Approach for Ensuring the Integrity of Private Databases

Asmaa Alqassab and Mafaz Alanezi

Abstract Watermarking implies embedding of data in a manner thatit is easily accessible by a user that's authenticated rather than other users, bearing in mind that the underlying data may experience some changes as a result of such data embedding. Reversible watermarking has been developed in advance of watermarking, whereby both of data quality and data recovery are guaranteed. For long time, the watermarking scheme to meet image, audio and video data type security was considered ever popular, but at the same time, since last several years ago relational database watermarking was really in mind and under study, even various watermarking approaches have been suggested. Nonetheless, these techniques yet are not sturdy enough against malignant attacks that can result in false insertion, modification or elimination leading to data performance and quality deterioration. So, and with aim of advancing and improving existing watermarking techniques, a new robust and reversible relational database watermarking method is suggested by which better outcomes were achieved.

Keywords Watermarking · Database integrity · Relational database · Reversible watermarking · RRW

1 Introduction

In this day and age, for research and organization, data dissemination across online is a significant job that comprises databases selling and buying as well. Data preservation has turned essential at current information era where each organization has to preserve bulks of private data, but by virtue of the technologies and Internet fast expansion, getting access to the digitized data has become extremely easy, Accordingly, it is

A. Alqassab (✉) · M. Alanezi
Department of Computer Science, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq
e-mail: asmaa_mow@uomosul.edu.iq

M. Alanezi
e-mail: mafazmhalanezi@uomosul.edu.iq

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_49

susceptible to security peril regarding data modification and ownership rights. Ergo, the necessity for copyright protection and ownership proving of databases arose. Cryptographic methods generally do not relate the cryptographic data with the digitized contents, and this is why the information of the ownership cannot be provided by cryptography, besides, no guaranteed procedure to track the digitized contents modification and redistribution is provided by the cryptosystem. So, cryptography solely will not be quite suitable for the protection of digital rights [1–7].

On another note, digital watermarking is one of the in demand and widely utilized techniques by which for a diverse data format, security is ensured in terms of tamper proofing and ownership protection. Watermarking techniques primely are employed to preserve data that's accessible to the public from being tampered, protect ownership of such data, secure integrity and similar other objectives, by providing appropriate watermarking techniques, one can ensure data protection and can tell whether or not any tamper has been made to the received data. Howbeit, watermarking may cause the precise recovery of cover work infeasible due to the loss of the sensitive information existing in the cover data. Embedding a watermark in a relational database without affecting its integrity is extremely hard, as the databases unlike the multimedia data, have too limited redundancy, that's to say, the database is a collection of autonomous objects of attributes or tuples then the watermark has to be inserted into these tuples, while multimedia objects comprised of great number of bits offering wide cover for watermark embedding. Watermarking methods have two fundamental sorts: fragile and robust watermarking. In fragile methods of watermarking, for tamper detection, a fragile watermark is utilized to specify any potential area where somebody has manipulated the watermarked data purposely or accidentally, this sort of watermarking possibly ruined or damaged after processing is implemented on the data content, while robust methods of watermarking are characterized with robustness that makes the watermarked data resistant against a diversity of malign attacks and benign alterations [8–13].

There are some cases that may demand applying the recovery process on the watermarked database to regain the original database, for instance, when database operations like delete, update, etc., is done by somebody on an original database, so the capability to do such actions is known as "reversibility" property in which the original database can be regained together with the embedded watermark. Withal, there is a possibility to have some data lost as a result of the watermark embedding in the database, but such problems do not happen with the reversible or lossless methods which insure original database recovery from the watermarked database. Conversely, considering about irreversible watermarking methods, they do not allow the recovery capability, yet the embedded watermark only can be regained from the watermarked database [14–16].

To solve above problems, we are introducing our suggested robust and reversible watermarking technique for relational databases.

The rest of the paper is structured into different sections as follows: an inclusive survey on this subject is presented in Related Works section, the new technique suggested for watermark embedding and extracting is described in Proposed Technique section, The Experimental Results section shows the analysis and results of the suggested method, and ultimately the last section finalizes the paper with summaries.

2 Related Works

So far, several robust reversible watermarking methods for relational database have been suggested. Varanasi A. et al. suggested an approach, in which in the preprocessing stage an appropriate watermarking function is chosen along with an optimization tech that's the genetic algorithm for an optimal watermark. Then, these parameters are utilized for both watermark insertion and extraction processes [17].

Jawad K. and Khan A. suggested an approach on the basis of the difference expansion but uses the genetic algorithm (GA) with the objective of distortion reducing and watermark capacity improving. For the watermark insertion process, in lieu of selecting fewer effective attributes, various attributes are selected using GA. With the GA fitness function, attribute-wise distortion and tuple are merged for the purpose of making it hard to figure out the watermarked attribute by any attacker [18].

Farfoura M. E. et al. suggested a reversible data embedding approach known as prediction-error expansion on integers to accomplish reversibility. In which even though 70% of the watermarked relation tuples were erased, it stills be possible to detect the watermark swimmingly. A secure key is used to embed the watermarks into the database relation. Additionally, a majority voting method is employed at the watermark extraction process to rectify bits of the watermark that has been extracted from the data [19].

Hu D. et al. suggested an approach for numerical data watermarking of the relational database, where a new histogram shifting of error prediction watermarking (HSW) and GA were amalgamated with the objective of robustness improving and distortion minimizing for the purpose of database watermarking. The suggested GAHSW approach ensures the watermarked database quality using HSW and ideally embeds the watermark data bits together in one group time and again using GA, furthermore, it has the capability to recover not less than half of the watermark data even when an attacker alters or removes till 90% of the tuples [20].

Li Y. et al. suggested a reversible database watermarking approach with low distortion on the basis of histogram gap owing to the wide gap in histogram of the database integer data type. Firstly, the suggested approach computes the tolerance of the attribute column which includes the whole integer data, and based on this tolerance it acquires the prediction error. Then, the tuples of the database are sorted haphazardly and the histogram is created using the acquired prediction error, and this can be done depending on the watermark data bits that will be embedded. Finally, to obtain the peak bin of the histogram, the correction rule of the histogram is utilized, then the amount of successive non-zero prediction errors on both right and left peak sides is found, and with a less amount of non-zero prediction errors, the histogram shift is carried out on the side to realize the embedding of the watermark [21].

Li Y. et al. suggested an approach premised on histogram continuous columns, in which and according to the grouping key and the length of the watermark, database tuples are sorted. Next, the histogram is created using the calculated prediction-errors absolute values. Then, to locate the whole continuous columns, the histogram needs to be traversed, the sum of the height of every continuous column is calculated and the set of the continuous columns possesses the greatest sum is chosen as the locations where the watermarks will be embedded [22].

Harish R. and Vijay K. S. suggested a reversible data inserting system named prediction-blunder expansion (PE), in which once the admin assigns a specific primary image for a user, he may login and be able to transmit covert images to other previously registered users. The covert image watermarked with the primary image through applying the pixel selection algorithm, and then encryption is applied on the watermarked covert image, and the resulted image is finally ready to be sent. Only recipient user has the capability to extract the watermarked image, any attacker may solely obtain the primary image but not the covert image [23].

Tale P. G. et al. suggested an approach that applies a reversible information embedding technique named prediction-error expansion on numerical data. Additionally, the suggested approach applies word shifting as feasible technique for data hiding in multi-word, non-numerical data type attributes of the chosen tuples. Using a secret key, the suggested approach, picks the tuples that will be marked along with the attributes that can hold the marks, thereafter, the mark will be embedded through horizontally shifting the position of a word belong to an attribute. To select which space or position will convey the mark, Levenshtein distance within two consecutive words of an attribute is utilized [24].

Parameswari M. et al. suggested an approach that embeds each bit of a watermark that created from Coordinated Universal Time date-time (UTC), in every chosen row of a numeric attribute. With the intention of embedding the watermark bits into a data set belong to a database owner, the watermark embedding procedure uses the bits of the watermark along with a secret key as input and as a result, the data set is transformed into a watermarked data set. For the purpose of choosing just those tuples possessing even hash values, MD5 cryptographic hash function is utilized on the chosen data set [25].

Hou R. and Xian H. suggested a graded reversible watermarking approach called (GRW), in which four basic algorithms were developed with the intention of facilitating the operations of each of: (watermark embedding, data quality grade detection, watermark detection and data quality grade enhancement). Prior to data spreading and to embed the watermark bits into the data, the data legitimate owner applies the (watermark partition embedding algorithm) by which data are split into a number of data segments, whereat every segment is considered as an autonomous unit for watermarking process, accordingly in any partition of the data, the watermark may be used for copyright claiming purpose [26].

3 Proposed Technique

RRW watermarking method is used so as to improve the data recovery process along with a robust security technique for the relational database. The proposed method includes four phases as follows:

1. preprocessing phase.
2. watermark embedding phase.
3. watermark extraction phase.
4. data recovery phase.

During the preprocessing phase, the user selects and prepares the image to be watermarked along with two types of watermarks: visible as logo and invisible which its data can be taken from about more than six of the relational database records together that has been created using various types of information samples (integer, real, character and date), after that, during the invisible watermark embedding phase, a new method based on DE is applied on the specified areas of the selected image to embed the watermark data in such a way that data of both of the image and the watermark will not be affected. During the watermark extraction phase, the watermark data is extracted without any changes. Finally, during the data recovery phase, the extracted watermark data is compared with the original watermark data. A proposed system flowchart can be seen in Fig. 1.

4 Experimental Results

In this section, experimental results are reported. The major purpose of such experiments is to test the robustness of the proposed method against unauthorized users and evaluate the data integrity within the proposed scheme.

To experiment and evaluate the suggested watermarking data and verification technique, we tested it in a machine having configurations Intel (R) Core i5 processor, 2.50 GHz, 8 GB RAM, and the operation system platform is Microsoft Wnidow10 feature experience pack.

Experimental results prove that the suggested method can maintain the integrity of the database data by providing the required security level.

The experimental results for the proposed RRW method with respect to both of PSNR and MSE values are obtained between both of the original and watermarked image are shown in Table 1.

The mentioned experimental results in table above show good results for both PSNR and MSE values after the watermark embedding process in five selected sample images with different sizes, which prove that the proposed method is effective and has no effect on the selected images no matter its size.

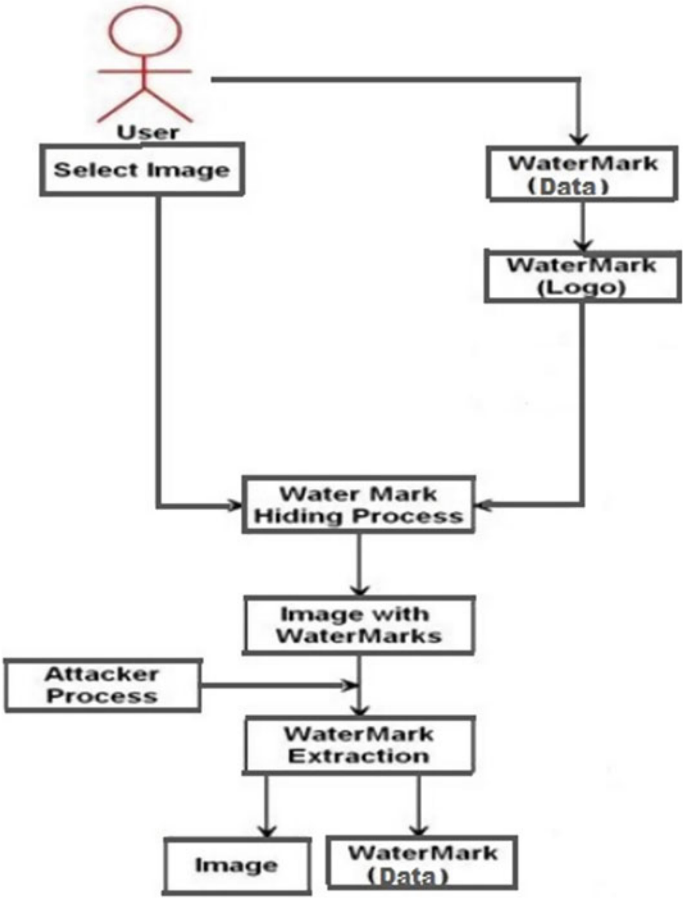


Fig. 1 Proposed system flowchart

5 Conclusion

Through the last few years, reversible watermarking researches on relational databases had been suggested to ensure both of information integrity and ownership protection. Principally, such schemes ensure original information recovery from the watermarked information, while irreversible watermarking techniques only provide ownership protection. Having such characteristic within the reversible watermarking methods allowed it to arise as a candidate solution for the information ownership protection purposes.

This paper suggested a novel reversible and powerful method for relational database watermarking. we have illustrated through the experimental results how this method can be applied to efficiently protect the database content integrity.

Table 1 Experimental results for the proposed RRW method

Image sample	Image size	PSNR	MSE
	190 * 266	78.35341727828342	0.0009500328731097962
	225 * 225	78.10788564088325	0.0010052910052910052
	198 * 255	79.95653747451512	0.0006567901234567901

(continued)

Table 1 (continued)

Image sample	Image size	PSNR	MSE
	275 * 183	80.74182409631939	0.0005481481481481482
	251 * 201	80.43958427026068	0.0005876543209876543

The suggested method is independent from the way by which database information is structured and how the records are organized. Additionally, the existence of reversibility property made it possible to recover the watermark along with the original data.

Acknowledgements The researchers thank the Department of Computer Science, College of Computer Science and Mathematics, University of Mosul.

References

1. Gadiya PR, Kale PA (2016) Reversible watermarking for relational data a brief review. Int J Comput Technol Appl (IJCTA) 7(1):25–28
2. Melkundi S, Chandankhede C (2015) A robust technique for relational database watermarking and verification. In: 2015 International conference on communication, information & computing technology (ICCICT), Mumbai, India

3. Churi AA, Shinde VD (2018) A review on reversible watermarking techniques. Int J Res Eng Appl Manage (IJREAM) 4(4):401–406
4. Al-Qassab AM, Nori AS (2014) Steganographic Technique using Fractal Image. International J Inf Technol Bus Manage (JITBM & ARF) 23(1):52–59
5. Kamran M, Farooq M (2018) A comprehensive survey of watermarking relational databases research. arXiv:1801.08271v1[cs.CR]
6. Iftikhar S, Kamran M, Anwar Z (2015) A survey on reversible watermarking techniques for relational databases. Secur Commun Networks, pp 2580–2603. https://doi.org/10.1002/sec.1196
7. Alanezi M, Altaay ISM, Malla'aloo SYH (2021) Payload and quality augmentation using steganographic optimization technique based on edge detection. Indones J Electr Eng Comput Sci 23(1):529–539. https://doi.org/10.11591/ijeecs.v23.i1
8. Kulkarni RD, Patil DV (2015) Watermarking of relational databases Survey. Int Res J Eng Technol (IRJET) 2(9):785–791
9. Khan A, Siddiqa A, Munib S, Malik SA (2014) A recent survey of reversible watermarking techniques. Elsevier, pp 251–272 (Information Sciences 279 (2014))
10. Murugan R, Abraham JT, Salim I (2019) A robust watermarking technique for copyright protection for relational databases. Int J Recent Technol Eng (IJRTE) 8(3):4040–4046
11. Dwivedi AK, Sharma BK, Vyas AK (2014) Watermarking techniques for ownership protection of relational databases. Int J Emerg Technol Adv Eng, pp 368–375 (International Conference on Advanced Developments in Engineering and Technology (ICADET-14), India)
12. Alqassab A, Alanezi M (2021) Relational Database Watermarking Techniques A Survey. J Phy Conf Ser pp 1–9
13. Murugan R, Jaseena KU, Abraham JT (2017) An invisible watermarking technique for integrity and right protection of relational databases. Int J Appl Eng Res 12(24):15754–15758
14. Chathuranga KMM (2019) Watermarking technology for copyright protection of relational databases. https://doi.org/10.13140/RG.2.2.21623.06565
15. Alfagi AS, Manaf AA, Hamida BA, Hamza MG (2014) A systematic literature review on necessity, challenges, applications and attacks of watermarking relational database. J Telecommun Electron Comput Eng 9(1–3):101–108
16. Shaikh UA, Shedge KN (2016) A secure watermarking technique for numeric and non-numeric relational data. Int J Innov Res Comput Commun Eng (IJIRCCE) 4(10):18013–18018
17. Varanasi A, Ravi G, Veerender D (2017) RRW—a ROBUST and reversible watermarking technique for relational data. Int J Professional Eng Stud (IJPRES) 8(4):90–93
18. Jawad K, Khan A (2013) Genetic algorithm and difference expansion based reversible watermarking for relational databases. J Syst Software 86(11):1–12
19. Farfoura ME, Horng S, Wang X (2013) A novel blind reversible method for watermarking relational databases. J Chin Inst Eng 36(1):87–97
20. Hu D, Zhao D, Zheng S (2018) A new robust approach for reversible database watermarking with distortion control. IEEE Trans Knowl Data Eng, pp 1–14
21. Li Y, Wang J, Ge S, Luo X, Wang B (2019) A reversible database watermarking method with low distortion. Math Biosci Eng 16(5):4053–4068
22. Li Y, Wang J, Jia H (2020) A robust and reversible watermarking algorithm for a relational database based on continuous columns in histogram. Mathematics, pp 1–12 (MDPI, Basel, Switzerland)
23. Harish R, Vijay KS (2017) A robust and reversible watermarking technique for relational data. Int J Adv Res Comput Commun Eng (IJARCCE) 6(5):57–61
24. Tale PG, Dharaskar RV, Thakare VM (2017) Reversible relational database watermarking using Levenshtein distance and prediction-error expansion. Int J Comput Math Sci (IJCMS) 6(5):27–34

25. Parameswari M, Vidhya M, Sharmila V, Sharmila JJ (2016) RWTT-reversible watermarking technique based on time-stamping in relational database. Middle-East J Sci Res 24(4):1517–1523
26. Hou R, Xian H (2019) A graded reversible watermarking scheme for relational data. Mobile networks and applications. Springer Science+Business Media, LLC, part of Springer Nature

Hybrid Optimization Approach for Adaptive Beamforming in Smart Antenna System

S. Samal, H. K. Sahoo, and Pradipta Kumar Das

Abstract The paper presents novel the adaptive beamforming in linear antenna array through hybrid optimization algorithm. The hybrid optimization algorithm has been formulated by combining improved version of the whale optimization algorithm (IWOA) and improved version of sine–cosine algorithm. Application of hybrid algorithm in the beamforming is to estimate the excitation weights of the desire signal applied on array elements, different interferences received from different directions and update the position the receiver so as to receive better quality of service by adopting the weight of the input signal. The robustness of the algorithm is confirmed through the simulation in MATLAB, and result shows that performance has been improved by minimizing bit rate, power transmitted, beamforming through different number of array elements using hybrid optimization algorithm.

Keywords Beamforming · IWOA · SCA · Smart antenna · Bitrate · Power transmitted

1 Introduction

In the present day, all over the world, the systems and technologies related to communication are changing rapidly. To enhance the performance of the systems new technologies are developing day-by-day. The next-generation wireless systems are required to have a high voice quality compare to the present cellular mobile standard and provide a higher bit rate. At the same time, the system should be operate in any type of surroundings (rural, urban, suburban), indoor and outdoor

S. Samal (✉) · H. K. Sahoo
Department of Electronic and Telecomunication Engineering, VSSUT, Burla, Odisha, India
e-mail: subhashreesmal12r@gmail.com

H. K. Sahoo
e-mail: harish_etc@vssut.ac.in

P. Kumar Das
Department of Information Technology, VSSUT, Burla, Odisha, India

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_50

environment, as well. In wireless transmission, the signal propagates through multipath and reached at the receiver at different time, delay, and amplitude. This cause of multipath fading increasing the quality or reducing the effective error rate in a multipath fading channel is extremely difficult, and increase of user in the wireless communication system leads to increase of interference in the signal. This difficulty can be minimized by use of adaptive antenna array [1]. The adaptive antenna array used the adaptive beamforming strategy to steer the radiation pattern by engaging the foremost beam toward desire signal and employing null toward to the interference which produce the higher signal to interference with noise ratio (SINR). The appropriate weight of the individual array elements determined the placement main beam toward desire signal and null toward undesired signal. The appropriate weight of individual array elements for adaptive beamforming application can be determined through optimization algorithm. Beamforming in smart antenna array has been resolved through many heuristic and meta-heuristic algorithms by the researchers. Linearly constraint minimum variance-based technique has been adopted for beamforming [2]. Partition-based particle swarm optimization (PSO) has been taken care to form optimal beamforming by minimizing square error [3]. Improved binary PSO has been adopted for optimal beamforming of planar antenna array [4]. Cultural algorithm has been exercised for beamforming in uniform circular array (UCA) [5]. A hybrid algorithm PSOGSA has been introduced for adaptive beamforming in smart antenna array and generate normalized array factor [6]. The optimization algorithm has been introduced for adaptive beamforming in smart antenna array [2–4]. Improved group-based PSO has been adapted for joint adaptive beam forming in wave communication [5]. Deeping learning algorithm has been illustrated for adaptive beamforming [7]. These algorithms provide weight of array elements to generate maximum SINR value. In some cases, the optimization algorithm could not able provide the better weight adjustment of antenna elements which leads to unacceptable radiation patterns and generates a lower SINR value. Therefore, it is essential to adopt the better optimization algorithm which avoids the premature convergence, trapped at local optimum and better weight adjustment in smart antenna array. The objective of the paper is to generate an effective and efficient position of the receiver to receive the better quality of service by avoiding different source of interference from different angles and maximize the radiation beam pattern toward the envisioned user or signal of interest. At the same time, minimize the radiation pattern toward the signal of not concentration through controlling of weight of uniform linear array (ULA) or UCA. The proposed work has been carried out through the hybridization of the optimization algorithm. The hybridization algorithm has been carried out by taking care of improved WOA and improved SCA. Sine–cosine algorithm is improved position-updated policy of the agents by using oppositional-based learning strategy, and at the same time, the WOA updates the position of search agent through cosine function, and best position of the search space is updated by comparing best position of the WOA and SCA. The aim of the hybridization is to maintain the better balance between exploration and exploitation, avoid stagnating in local optimum, and improve in the convergence toward global optima.

The rest of the paper is illustrating as follows: The design of the fitness function for beamforming in smart antenna array is elaborated in Sect. 2. Improved whale optimization algorithm is described in Sect. 3. Section 4 described the improved sine–cosine algorithm through oppositional-based learning strategy and also algorithm for beamforming in smart antenna array. Simulation result and its analysis have been illustrated in Sect. 5. Conclusion of the work is described in Sect. 6.

2 Paper Preparation Design of Fitness Function for Beamforming in Smart Antenna System

Consider a uniform linear antenna array (ULA) of M-element employed for adaptive beamforming receiver that receives a desired signal $s(k)$ and N interferences signals on the array at ith time stamp is expressed as:

$$\begin{aligned} X(t) &= s(t)a(\theta_d) + \sum_{i=1}^{N} i_i(t)a(\theta_i) + n(t) \\ &= d(t) + u(t) \end{aligned} \tag{1}$$

where $s(t)$ and $i_i(t)$ are the desire signal and ith inference at time stamp t, respectively. $d(t) = s(t)a(\theta_d)$ denotes the desire input signals; $u(t) = \sum_{i=1}^{N} i_i(t)a(\theta_i) + n(t)$ denotes the inference input signal. $n(t)$ represents the vector of M uncorrelated noise signal. θ_i and θ_d signify the direction of interference ith and desire signal, respectively. $a(\theta_i)$ and $a(\theta_d)$ signify the steering vectors of $i_i(t)$ and $s(t)$, respectively, and it is presented as follows:

$$a(\theta_i) = \left[1,\ e^{-j\frac{2\pi}{\lambda}q\cos\theta_i}, \ldots, e^{-j(M-1)\frac{2\pi}{\lambda}q\cos\theta_i}\right]^T \tag{2}$$

$$a(\theta_d) = \left[1,\ e^{-j\frac{2\pi}{\lambda}q\cos\theta_d}, \ldots, e^{-j(M-1)\frac{2\pi}{\lambda}q\cos\theta_d}\right]^T \tag{3}$$

where q represents the space between adjacent element of ULA and its value is determined as:

$$q = \frac{\lambda}{2} \tag{4}$$

λ denotes the wavelength of carrier signals, and T is the transpose operation. The output of the beam former array is illustrated as follows:

$$y(t) = w^{\mathrm{H}} X(t) \tag{5}$$

where $w = [w_1\, w_2\ w_3, \ldots, w_M\,]^{\mathrm{T}}$ denotes weight of the beamforming smart antenna and H denotes Hermitian transpose. The output array for the desire signal power is illustrated as:

$$\alpha_d^2 = E\Big[\big|w^{\mathrm{T}} d(t)\big|^2\Big] = E\Big[\big|w^{\mathrm{T}} s(t) a(\theta_d)\big|^2\Big] \tag{6}$$

The output array power of the interference is illustrated as:

$$\begin{aligned}\alpha_u^2 = E\Big[\big|w^{\mathrm{T}} u(t)\big|^2\Big] &= E\left[\left|w^{\mathrm{T}}\left[\sum_{i=1}^{N} i_i(t) a(\theta_i) + n(t)\right]\right|^2\right] \\ &= w^T \sum_{i=1}^{N} R_i w + w^{\mathrm{T}} R_n w \end{aligned} \tag{7}$$

where E denotes the expectation operator. $R_i = E\big[i(t) i^{\mathrm{T}}(t)\big]$ signifies correlational matrix of the ith interference, and $R_n = E\big[n(t) n^{\mathrm{T}}(t)\big]$ is noise correlational matrix. The fitness function is designed based on inverse of the signal to inference plus noise ratio which is presented as:

$$F = \frac{\alpha_u^2}{\alpha_d^2} \tag{8}$$

The fitness function is minimized to steer the peak of the main lobe toward the signal of interest and generates null patterns in the direction of arrival of all undesired signals. The noise variance is computed through the signal-to-noise ratio and presented as:

$$\alpha_{\text{noise}}^2 = 10^{-\mathrm{SNR}/10} \tag{9}$$

3 Improved Whale Optimization Algorithm

Whale optimization algorithm is a population-based meta-heuristic algorithm for generating global optimal solution of the optimization problem in the search space. It is mimicking the behavior of humpback whales to get their food. It generates the optimal solution in three processes: search for prey, encircling prey, and bubble net attacking. The exploration process of the optimization algorithm is achieved through the search for prey and encircling prey, and at the same time, the exploitation process is achieved through the bubble net attacking. During the searching for prey, the search agent follows another search agent in the population, and position of the search agent is expresses as follows:

$$X(t) = X_{\text{rand}} - P.G_{\text{rand}} \tag{10}$$

$$G_{\text{rand}} = |A.X_{\text{rand}} - X(t)| \tag{11}$$

where P and A are the coefficient factor, X_{rand} is the position vector of the search agent selected randomly from the current population in the search space. t is the current iteration. The value of P and A is evaluated through the following expression as follows:

$$\begin{aligned} P &= 2.b.r - b \\ A &= 2r \end{aligned} \tag{12}$$

where r is the random number in the range of 0 and 1. The value of b is decreased from 2 to 0 with iterations. During the encircling prey, the search agent follows the best optimal solution found so far and updates the position based on the target position.

$$G_{\text{rand}} = \left|A.X^{\text{T}}(t) - X(t)\right| \tag{13}$$

$$X(t+1) = X^{\text{T}}(t) - P.G_{\text{rand}} \tag{14}$$

where $X^{\text{T}}(t)$ is the position vector of the target. The spiral model is considered for movement of the search agent toward the target, and it is expressed as:

$$X(t+1) = G^{*}.e^{bl}.\cos(2\pi l) + X^{\text{T}}(t) \tag{15}$$

$$G^{*} = \left|X^{\text{T}}(t) - X(t)\right| \tag{16}$$

where G^{*} is the measure distance between the search agent and target, l is the random number within range of -1 and 1, and b is the constant number and considered as 1. The value of l is expressed as follows:

$$\begin{aligned} l &= (a_1 - 1) * \text{rand} + 1 \\ a_1 &= -1 - \frac{t}{\text{max \_iteration}} \end{aligned} \tag{17}$$

The search agent updates the position their position based on the probability value and expressed as below:

$$X(t+1) = \begin{cases} X^T(t) - P.G_{\text{rand}} & \text{if } s \prec 0.5 \\ G^*.e^{bl}.\cos(2\pi l) + X^T(t) & \text{if } s \geq 0.5 \end{cases} \tag{18}$$

where s is the random number in the range of 0 and 1. The cosine function is used in WOA to modify the position of the search agent and expressed as below:

$$X(t+1) = \begin{cases} X^T(t) - \cos(2\pi l).G_{\text{new}} & \text{if } s \prec 0.5 \\ G^*.e^{bl}.\cos(2\pi l) + X^T(t) & \text{if } s \geq 0.5 \end{cases} \tag{19}$$

$$G_{\text{new}} = A.X^{\mathrm{T}}(t) - X(t) \tag{20}$$

4 Sine–Cosine Algorithm (SCA) Improved Through Oppositional-Based Learning

It is a population-based meta-heuristic algorithm which is used to update the position of the agent in the search space through the function of sine and cosine. The position of the agent is updated based on the following expression.

$$X_i(t+1) = \begin{cases} X_i(t) + p_1 \times \sin(p_2) \times \left|p_3 P_i^g - X_i(t)\right|, & p_4 \prec 0.5 \\ X_i(t) + p_1 \times \cos(p_2) \times \left|p_3 P_i^g - X_i(t)\right| & p_4 \geq 0.5 \end{cases} \tag{21}$$

where p_1 and p_2 select the direction of movement and amplitude of the movement in the search space, respectively. p_3 is a random weight which regulates the effect of the goal as the iteration continues, and p_4 is the control parameter for adopting sine or cosine function to update the position of the particles in the search space, and the value of the control variable lies within [0,1]. The oppositional-based learning strategy has been adopted to select the best position of the particle in the region of the search. The opposite real number $X \in [l, u]$ is expressed by $\overline{X}$.

$$\overline{X} = u + l - X \tag{22}$$

where u and l signify the upper and lower bound of the search space, respectively. In the multidimensional search space, the agent position X can be expressed as follows:

Consider $X = [x_1, x_2, x_3, \ldots, x_n] \in R^n$, where $x_1, x_2, x_3, \ldots, x_n \in R$ and $x_i \in [l_i, u_i]$. The opposite point $\overline{X} = [\overline{x}_1, \overline{x}_2, \overline{x}_3, \ldots, \overline{x}_n]$ is expressed by:

$$\overline{x}_i = u_i + l_i - x_i \quad i = 1,\ 2,\ 3, \ldots, n \tag{23}$$

The best solution is obtained by taking the union of the current population and opposite population of current one. It is expressed as follows:

$$X_{\text{new}} = X \cup \overline{X} \tag{24}$$

$$X_{\text{best}}(t) = \min(f(X_{\text{new}}(t))) \tag{25}$$

The position of the agent is updated through the best position $X_{\text{best}}(t)$ obtained through oppositional-based learning and expressed as follow:

$$X_i(t+1) = \begin{cases} X_{\text{best}}(t) + p_1 \times \sin(p_2) \times |p_3 X_{\text{best}}(t) - X_i(t)|\ , & p_4 \prec 0.5 \\ X_{\text{best}}(t) + p_1 \times \cos(p_2) \times |p_3 X_{\text{best}}(t) - X_i(t)| & p_4 \geq 0.5 \end{cases} \tag{26}$$

The parameter p_3 is the random number lies in between [0,1], and p_1 and p_4 are modified as follows:

$$p_1 = \mu\left(1 - \frac{t}{\max\_\text{iteration}}\right) \tag{27}$$

where μ positive constant and its value is set as 0.7

$$p_4 = 1 - \frac{t}{\max\_\text{iteration}} \tag{28}$$

Each individual algorithm exhibits to be trapped at local optima in the large search space. Therefore, it is necessary to hybridized improved WOA, and improved version of SCA has been hybridized for improving the balance between exploration and exploitation capability, avoid stagnating at local optimum, improve the convergence speed as individual algorithm unable produce better quality of solution in the search space. The IWOA–SCA has been integrated to perform the optimal beamforming in smart antenna array by removing the undesired signal and update the better position of the receiver to receive the better quality of service. The algorithm used for beamforming in smart antenna is elaborated as follows:

Algorithm 1: Pseudocode for IWOA-SCA for beamforming in smart antenna array

Input: N is the Population, max_iteration is the maximum allowable iterations and initialize the population $X_i (1 \leq i \leq N)$. Initialize the parameters of MWOA and SCA

Output: Optimal Beamforming

1. Initialize the population
2. While($t < \max\_iteration$)
3. Evaluate the fitness value of X and identify best search agent $X^T(t) = \min(f(X))$
4. For $i = 1 \ to \ N$ (number of whales)
5.   Evaluate $b, s, a1, l$
6.   If (s<0.5)
7.   Compute G_{new}
8.   Update $X_i(t+1)$ using Eq.18
9.   else
10.   Update $X_i(t+1)$ using Eq.18
11.   If end
12. End for
13. $if(f(X_i(t+1) \prec f(X^T))$
14. $X^T(t) = X_i(t+1)$
15. For $i = 1 \ to \ N$
16.   If($p \prec 0.5$)
17.   Update the position $(X_i)_{SCA}$ through Eq 26.
18.   else
19.   Update the position $(X_i)_{SCA}$ through Eq 26.
20.   If($f(((X_i)_{IWOA}) \prec f((X_i)_{SCA}))$)
21. $X_i = (X_i)_{IWOA}$
22.   Else
23. $X_i = (X_i)_{SCA}$
24. for End
25. $t = t+1$
26. End while
27. Return global best solution

5 Simulation Result and Performance Analysis

The performance of the beamforming in smart antenna has exhibited through the simulation, and simulation has been carried out in MATLAB. The performance of the proposed algorithm has exercise to present bit error rate over different channels. The performance has been evaluated for different number of transmitting antenna, and its value is considered as $M = 4$, $M = 8$, $M = 16$. The outcomes of the result are presented in Figs. 1, 2, and 3 using proposed algorithm and its counterpart algorithm. The conclusion drawn from Figs. 1, 2, and 3 that bitrate is minimized with respect to the noise ratio for different value of channels in smart antenna array through the proposed algorithm as compare to its counterpart algorithms. Similarly, the performance has been evaluated for calculating the average power transmitted and average received SINR. The average transmitted power with respect to the number of receiver has been presented in Fig. 4 through different algorithms. The outcomes of Fig. 4 show that IWOA–SCA provides better beamforming at the receiver site with less

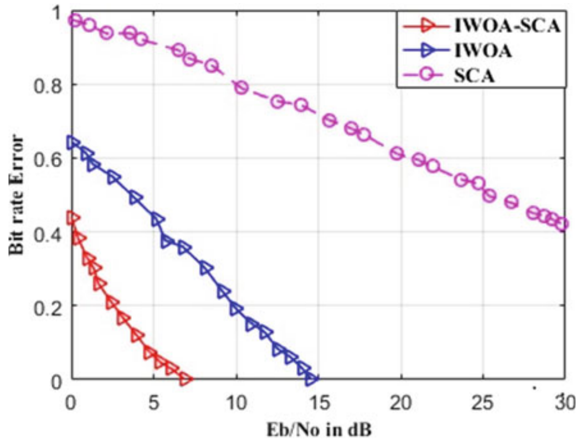


Fig. 1 Bit rate with noise ratio for $M = 4$ using proposed algorithm and its counterpart

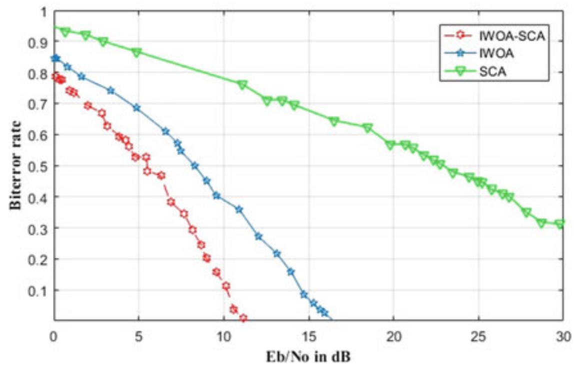


Fig. 2 Bit rate with noise ratio for $M = 4$ using proposed algorithm and its counterpart

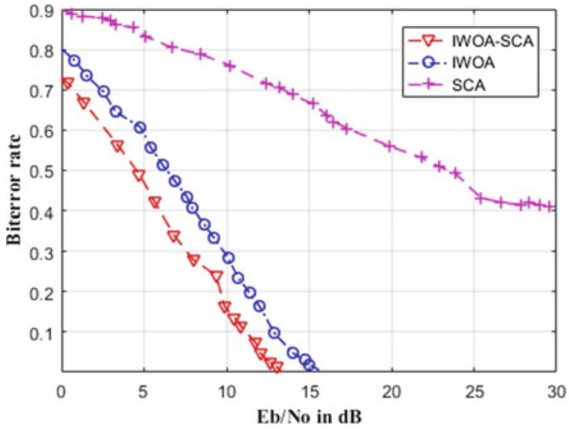


Fig. 3 Bit rate with noise ratio for $M = 16$ using IWOA and its competitor algorithm

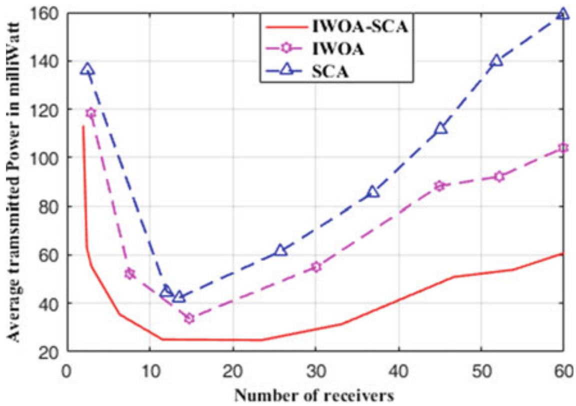


Fig. 4 Average transmitted power with number of receiver through different algorithms

transmitted less power as compare to the other algorithms. The average received SINR is plotted with number receiver through the different algorithms, and it has been illustrated in Fig. 5. The conclusion drawn from Fig. 5 that SINR is less in IWOA–SCA as compared to its competitor algorithms. Again, the performance has been evaluated in terms of average power transmitted with respect to number of iterations for beamforming through ULA and circular array elements, and it is elaborated through Fig. 6. Figure 6 shows that power transmitted is less in ULA than circular array elements and reduce the interference and enhance the system capability. The receive power with number of iteration has been elaborated in Fig. 7. Figure 7 shows that average power received is least in IWOA–SCA as compared to its counterpart algorithms.

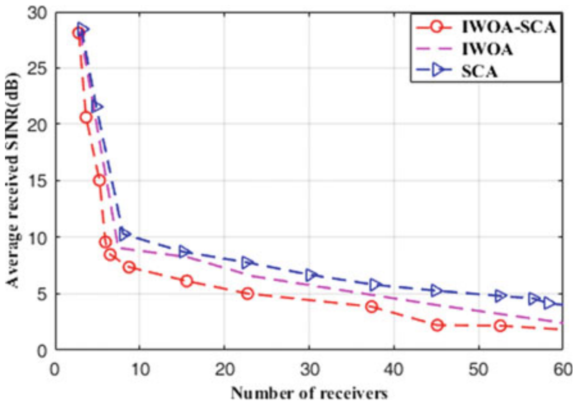


Fig. 5 Average received SINR with number of receiver through different algorithms

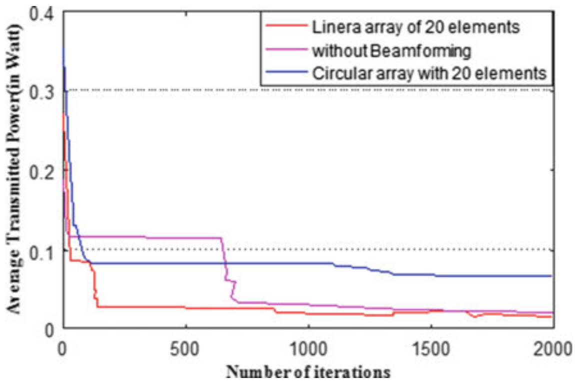


Fig. 6 Average transmitted power with number of iterations through for beamforming for 20 elements

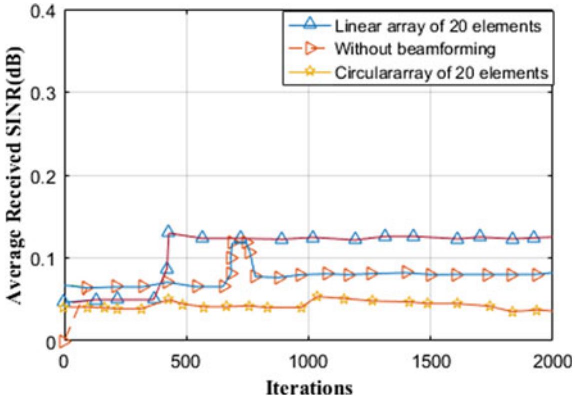


Fig. 7 Average receive power with number of iterations through for beamforming for 20 elements

6 Conclusions

The adaptive beamforming in smart antenna array has been carried out through the hybridization of optimization algorithm. The paper designed the fitness function by considering different constraints in radiation patterns and optimized through hybridization of IWOA and SCA algorithm to reduce the interferences and improve the quality of service in the terms of bitrate, power transmitted, and SINR. The simulation has been carried out through MATLAB to demonstrate the results, and results show that the proposed algorithm is outperformed over other counterparts algorithms.

References

1. Lazarus K, Noordin NH, Ibrahim Z (2018) An enhanced simulated Kalman filter algorithm and its application in adaptive beamforming. In: 2018 IEEE international RF and microwave conference (RFM). IEEE, New York, pp 321–324
2. Darzi S, Tiong SK, Tariqul Islam M, Rezai Soleymanpour H, Kibria S (2016) An experience oriented convergence improved gravitational search algorithm for minimum variance distortionless response beamforming optimum. PLoS One 11(7):1–20
3. Noordin NH, Arslan T, Flynn B, Erdogan AT (2013) Faceted array antennas for adaptive beamforming applications. In: IEEE international symposium on personal, indoor and mobile radio communications, PIMRC, pp 425–429
4. Darzi S, Sieh Kiong T, Tariqul Islam M, Rezai Soleymanpour H, Kibria S (2016) A memory-based gravitational search algorithm for enhancing minimum variance distortionless response beamforming. Appl Soft Comput J 47:103–118
5. Ramya P, Valarmathi RS, Poongodi C (2021) Antenna selection with Improved group based particle swarm optimization (IGPSO) and joint adaptive beam forming for wideband millimeter wave communication. J Ambient Intell Humaniz Comput 12(3):4291–4302
6. Magdy A, El-Ghandour OM, Hamed HFA (2015) Adaptive beam-forming optimization based hybrid PSOGSA algorithm for smart antennas systems. In: Progress in electromagnetics research symposium proceedings, Prague, Czech Republic, pp 973–977
7. Bianco S, Napoletano P, Raimondi A, Feo M, Petraglia G, Vinetti P (2020) AESA adaptive beamforming using deep learning. In: 2020 IEEE radar conference (RadarConf20). IEEE, New York, pp 1–6

IoT for Fight Against COVID-19

Kanta Prasad Sharma, Kirti Walia, and Shaurya Gupta

Abstract An emerging technology Internet of things is the backbone for better solution in medical science research, COVID infected sampling analysis, and device integration process. 2020 may be a year of healing not only for our mother earth, but for mankind too. It is a year of change and practice to develop ourselves against adversities. Due to the recent pandemic caused by COVID-19, many lives were affected. COVID-19 has created an adverse effect on the economy, education, mental health, and physical health of humans. It has been witnessed that despite lockdown, the death rate has increased. From several statistics, it can be observed that populations with less immunity have a higher mortality rate. This study has been performed to make a checkpoint on the factors which may be responsible for determining immunity level and based on that a prediction model may be prepared using a machine learning algorithm. This proposed work employs an IoT application to collect real-time symptom data from users to identify suspected coronavirus symptoms. IoT's sensor-based mechanism adopts for enhance capability of risk minimization specially in surgery cases compilation for COVID-19 type pandemic.

Keywords Adaptive immunity · Stress · RNA · SARS · Immunity · Comorbidity

K. Prasad Sharma (✉) · K. Walia
University Institute of Computing Chandigarh University Mohall, Mohali, India
e-mail: tokpsharma@gmail.com

K. Walia
e-mail: kirti.e8889@cumail.in

S. Gupta
Department of Computer Science and Engineering, UPES, Dehradoon, India
e-mail: shaurya55@gmail.com

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_51

1 Introduction

The technology is very important to fight against various medical viruses like corona, plugs, swine flu, etc., to secure human life. Internet of things is helpful for analysis and alerting humans regarding medical issues, so a human can immediate take action. Now, medical science rapidly adopts the technological support for reshaping modern healthcare centers, economical support.

The global challenges of the COVID-19 virus pandemic condition, these virus series generate global health crisis. The disease has found similar symptoms like fever; cough, and fatigue are recognized daily diagnosis process.

So, medical science needs an appropriate digital information system for COVID-19 pandemic, which control by IoT for overcome physical data analytic and crucial applications to fulfill higher efficiency requirements.

The current scenario of complete lockdown throughout the whole world has given Mother Nature time to heal and reincarnate itself into a new form, but human life has undergone various mental and physical illnesses. According to the statistics received by August 08, 2020, the present COVID patient count is nearly 1.93 million, and the total death count is 7.18 lacs. The whole world is slowly welcoming a new normal in all aspects. The perspective of the whole world has changed a lot. The recent statistics related to COVID-19 patients have proven that it is an RNA-based virus that has a high-contamination rate than other contagious diseases [1–7]. Though the vaccine is invented and many people are being immunized but still the possibility of attaining complete accomplishment benchmark against COVID-19 is under scrutiny. This paper focuses on the precise area of developing one's immunity level to win the battle again such kinds of diseases in a preventive manner, by enhancing the individual immunity level [8, 9].

1.1 Concept of Immunity

Immunity can be defined as the layer of protection of our body against infectious diseases, which in turn is quite capable to distinguish itself as foreign material and fighting against it whenever the body is under any infection attack [10]. The base mechanism is mainly concerned with disposing of foreign substances. The response of immunity mainly depends on the antigen–antibody response. Immunity response mainly depends upon two cells such as T cell and B cell [10].

1.2 Types of Immunity

Natural or Innate Immunity—It is an overall genetic characteristic like sex, age, overall lifestyle, etc. Due to the presence of this natural immunity, all persons have a certain level of immunity against certain diseases [10, 11]. With age, the potential of this immunity decreases. It may vary accordingly from an individual to another. Community transmission takes place if the immunity is weakened among individuals. Individual immunity depends upon certain characteristics like the health status of the person, nutritional habits, personal hygiene, or genetic habits.

Acquired Immunity—Specific immunity acquired through the development of antibodies inside human body [10]. Acquired immunity is immunity which is not natural. It is learned with time, and it starts developing as and when a person's immune system comes across foreign invaders then recognizes oneself substances which are called antigens. Therefore, the constituents of acquired immunity acquire the finest means to attack every antigen, and thereby, it begins to cultivate retention for that specific antigen. Acquired immunity is termed specific immunity, because it modifies its attack to a specific antigen that has been beforehand encountered. Its trademarks include the aptitude to acquire, familiarize recall.

1.2.1 Factors Accountable for Immunity Development

Certain habits may help in developing immunity within us. Those factors are as follows:

- **Hand Washing—**It is a good personal hygiene habit that helps to protect against pathogens. This practice boosts immunity by maintaining personal hygiene.
- **Sleep Cycle—**Sufficient sleep reduces the level of stress hormone cortical, which in turn increases natural immunity. Therefore, it is recommended to have a minimum of 6–8 h of sleep.
- **Nutritional Value—**Maintaining a balanced diet with nutritional value in our daily meal with enough amounts of minerals, vitamins, proteins, and carbohydrate is very much recommended.
- **Cortisol Level—**The stress level in our day-to-day life is another reason for abating our immunity. According to a report of the American Psychological Association, 75% of Americans experience medium to high-level stress [12].

The above factors are very important to affect the immunity of a person. Below is a figure to demonstrate different types of immunity (See Fig. 1).

1.3 World Statistics About Immunity Status

Though this research field is very novel, researches are effectively researching in this area. From some studies, it has been established that a relationship exists between the immunity and death rate by using the comorbidity factors. In the following figure (See Fig. 2), we can observe

Along with geographical area, the death rate in COVID-19 varies by age and gender also. In the following figure, this has been depicted (See Fig. 3) [13].

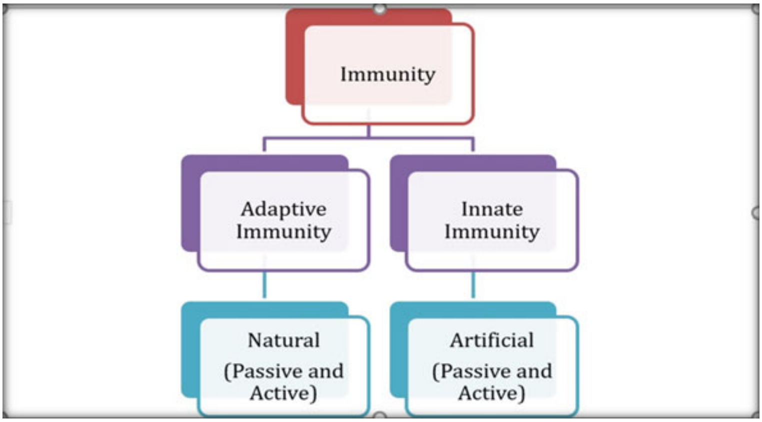


Fig. 1 Types of immunity

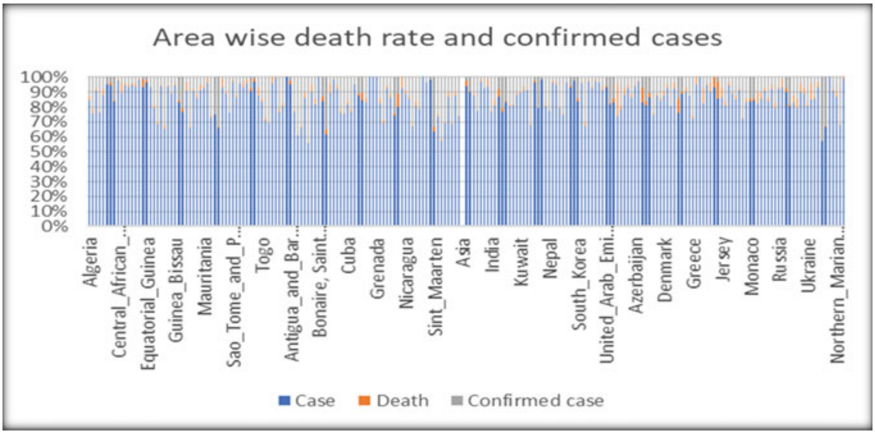


Fig. 2 Region wise death rate and confirmed cases for COVID-19

The comorbidity in COVID-19 is very important since it creates a lot of impact on the COVID-19 death rate. It has been observed that the patients with high comorbidity are having a high-death rate (Figs. 4 and 5).

Emerging technology—Internet of things-based applications provide high-quality outcomes, which becomes a new innovation and the best services especially for complicated cases handle and control during pandemic scenario [13]. So, IoT provides a new solution for medical science for doctors, surgeons, and patients, see in Fig. 6.

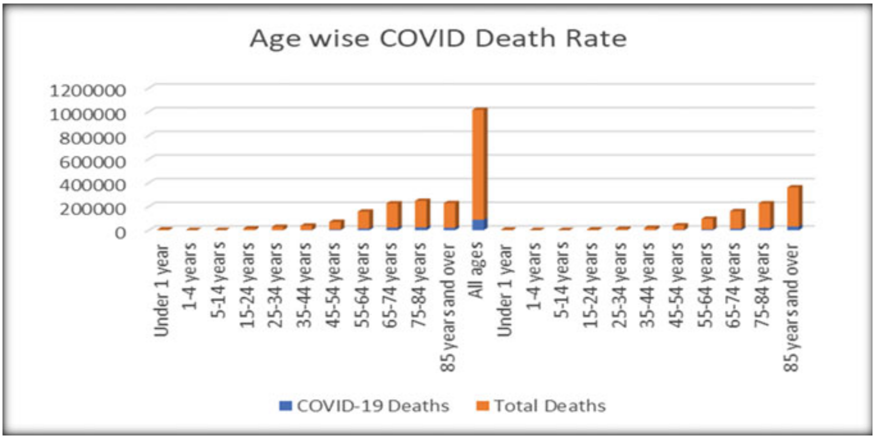


Fig. 3 Age-wise COVID-19 death rate

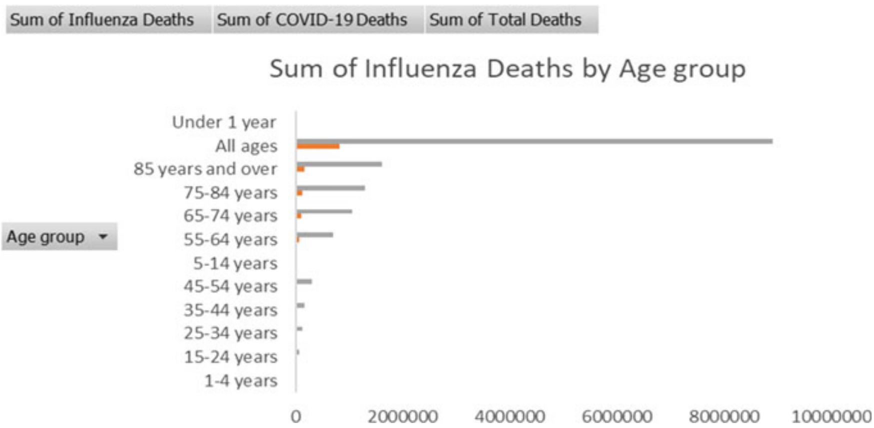


Fig. 4 Death in influenza [14]

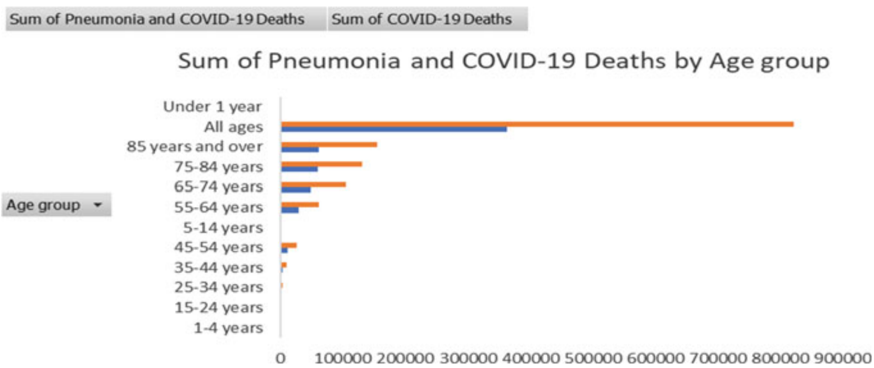


Fig. 5 COVID-19 deaths by influenza [14]

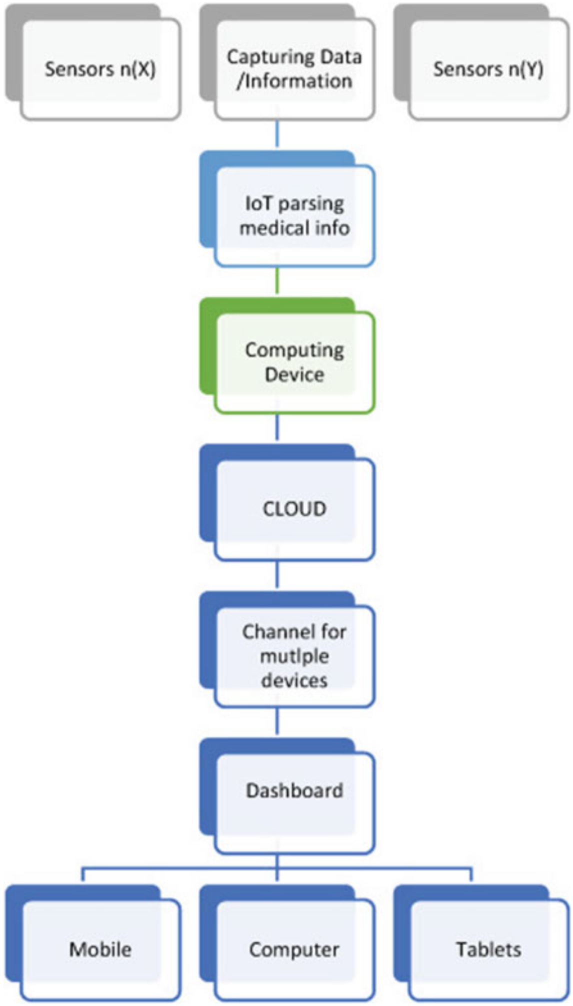


Fig. 6 IoT utilization in medical science

2 Related Study

The crucial situation of COVID-19 has changed the world's scenario, and from the very beginning, people are involved in research related to various aspects of COVID-19 which are affecting the whole world. One popular area in comorbidity field of patients. From some studies, it has been observed that the patients having diseases like diabetes, COPD, or cancer are in more risk zone. Certain statistics and reports have been published related to the comorbidity of COVID-19[1]. Scientists can establish a relationship with comorbidity [1–5]. Researchers have discussed several issues and factors related to immunity and how boosting immunity can help to protect mankind against the fight of COVID-19 ([3, 4, 15]). Along with other effects of

COVID on our daily lives, scientists are also concerned about our mental as well as physical well-being. From a survey, it has been observed that the death rate in COVID-19 is 1–3%, whereas the death rate in SARS or MERS is 10% and 35%, respectively [16]. According to the severity of death, corona-related death is not so severe but the contagiousness of the disease is much higher than others. The only key to protest against such highly infectious diseases. It has been observed by recent research, and those oncological patients are more prone to be affected in corona [6]. According to the study, some diseases like diabetes, hypertension, or cancer are very fatal for COVID-19. Another very interesting study performed by Soumier and Sirigu revealed that oxytocin secretion can actually boost our immunity power, but it is not a proofed study, and the research is still going on [7]. From another recent study, it has also been revealed by Carminati and his co-researchers that people who use BCG vaccine are immune to coronavirus more than others without the vaccine [16–18]. There are certain studies related to immunity aspect related to COVID-19. According to one study, it has been observed that many undocumented asymptotic symptoms may lead to herd immunity [19]. In their study, Kissler and his team have performed research on developing a mathematical model for assessing the lockdown period which is according to them is needed to maintain for the year 2022 [19]. The scientists have used deterministic model to get the assessment [20, 21]. From the study of Milo Schield, it is very much clear from a study performed by Milo Schield, that comorbidity has an impact upon coronavirus recovery rate [22]. It has been observed that 80% of patients have a mild or moderate disease and 13.8% having serious disease like high-blood pressure or any respiratory problem. According to the study of Scientist Roman Welfel and his team, coronavirus is an acute form of respiratory tract infection, and it solely depends upon individual immunity level, the severity of the infection varies from an individual to another individual [23]. By several studies, a conclusion can be done that immunity is a big factor to fight against highly contagious disease like corona and it is proofed that T cell which is responsible to fight against this pandemic [19, 24]. T cell is the cells responsible to identify and kill the pathogen and infected cells [21]. Most researchers in this field have focused on comorbidity like high-blood pressure, sugar, or obesity, and these are among the common comorbidity factors [25]. Among a lot of COVID cases, it has been observed by a study performed in New York that within a group of 5000+ where 57% having high-blood pressure and 41.7% having the problem of obesity [17]. So, from the studies of various researchers, one direction can be explored that immunity has a relationship with the capacity of protection against COVID-19.

3 Proposed Methodology

The proposed methodology has been designed on the basis of a questioner prepared and the statistical analysis based on received data. The statistical analysis has been performed on a set of data mostly based on primary data collected through a questioner and secondary data available in sites [1, 17].

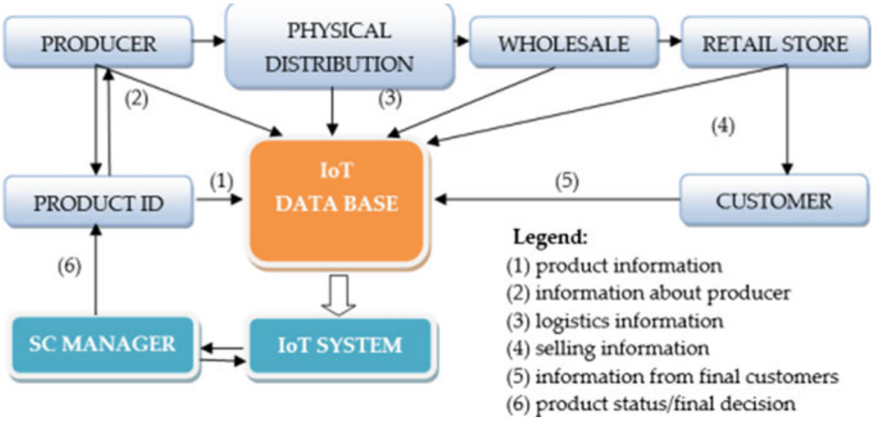


Fig. 7 IoT-based supply chain methodology

The proposed immunity data process using integrated IoT platform for supply chain management application, which are security, safety and sustainability, data portability, and predictability. IoT-based proposed method (seen Fig. 7) provides transparency and data visualization like pandemic result, stock of medical equipments, and real-time sample data collection and increases the transparency for physical distribution.

In light of ordinary and modern treatments of COVID-19 patients, who need a classical Ataxia-Telangiectasia trial to be better analysis of sampling dataset immunity level [13]. The main steps to prepare our proposed framework were as follows:

Step1: Prepare a questioner on immunity level measurement
Step 2: Distribution of questioner through different distribution channels
Step 3: Sample data collection
Step4: Statistical analysis of that data
Step5: Prepare a machine learning model for prediction of the effect
Step6: Result analysis.

4 Data Collection and Sampling

The result has been produced both from primary as well as secondary data. The prime source of primary data collection is questioner, whereas the secondary source of data is dataset available on different sites [17]. The primary source of data has been collected through survey done among xxx no. of Indian habitants, and the following result has been witnessed in Table 1.

Performance Evaluation: To evaluate the performance on the learning algorithms and process, that is based on accuracy, root mean square error, F-measure, and ROC area. These measures can be analysis based on confusion conditions like.

Table 1 Data analysis of survey performed considering immunity

Demographic data				
Gender	Male	Female	Transgender	Not prefer to say
Age group				
Area	Metro	Urban	Semi-urban	Rural
Questioner				
Is the patient diabetic?	Yes		No	
Has the patient high-blood pressure?				
Does the patient have any critical illness?	Yes		No	
Exercise habit				

False Positive (FP): We consider number predicative instance as positive, but they are actually negative.

False Negative (FN): The instances were classified predictive model as negative, but they are actually represented positive.

True Negative (TN): The number of instances which classified predictive instance as negative, and they are actually negative (Table 2).

The accuracy of a classifier is computed as the number of correctly classified instances to the total number of instances. It is given by (Table 3):

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

Table 2 Confusing conditions

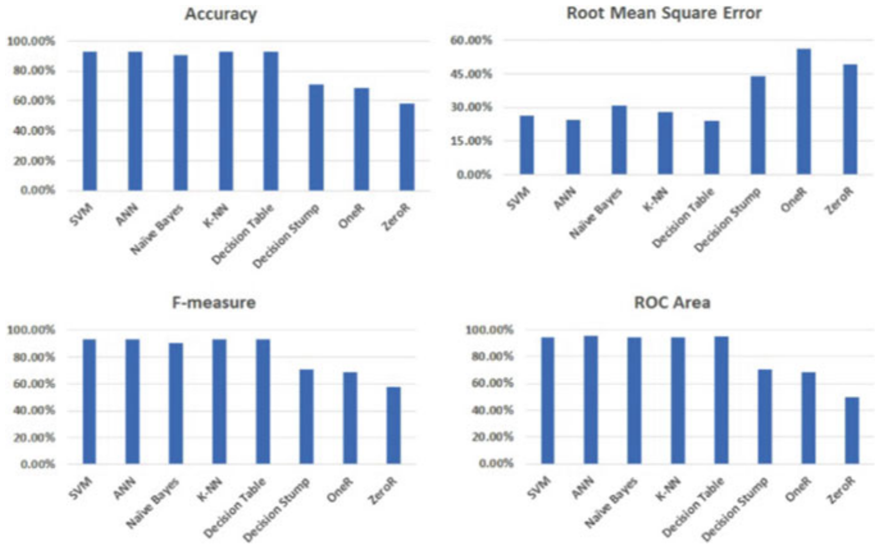
True positive (TP)	False negative (FN)
False positive (FP)	True negative (TN)

Table 3 Result performance

Techniques	Accuracy (%)	Root mean square error (%)	F-measure (%)	ROC area (%)
Support vector machine (SVM)	80.95	26.54	93.0	93.9
Neural network	80.89	24.23	92.9	95.5
Naïve Bayes	80.58	30.99	90.6	94.2
K-nearest neighbor (KNN)	80.89	28.06	92.9	93.9
Decision table	80.95	23.97	93.0	95.0
Decision stump	70.73	43.86	70.6	70.1
OneR	58.36	56.25	68.5	68.3
ZeroR	52.86	49.38	57.9	49.7

$$\text{True Rate} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

$$\text{False Rate} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{3}$$



5 Conclusion and Future Direction

This work reflects the relationship of immunity with COVID-19 disease and probability of protection from this highly contagious disease. IoT-based framework to reduce the impact of communicable diseases there are many researchers working in the comorbidity area and from their researchers a perfect conclusion can be drawn that by increasing immunity we may be able to fight against this new pandemic.

Emerging technologies can provide solution for manual record keeping. For a well-informed decision, IoT-based application will minimize the issues provide precise outcomes. By using IoT, modern healthcare equipments provide effective solution to save human life. In future, this technology will create innovation solution for the excellent surgery, sampling data analysis, and the best treatment of the patient to stay healthy and will used to any COVID-19 type pandemic.

References

1. https://www.worldometers.info/April (2020) 'Immunity passports' in the context of COVID-19. Available https://www.who.int/news-room/commentaries/detail/immunity-passports-in-the-context-of-covid-19
2. Hesary FT, Akbari H, The powerful immune system against powerful COVID-19: a hypothesis
3. Chowdhury MA et al (2020) Immune response in COVID 19: a review. J Inf Public Health
4. Jahandideh D, Taheri A (2020) Old medication for a novel disease? Med Hypothesis 12:109830
5. Cherri S, Noventa S, Zaniboni A (2020) Is the oncologic patient more susceptible to covid19 but perhaps less likely to undergo severe infection-related complications due to fewer cytokines storm as a consequent of the associated immunodeficiency? Med Hypotheses 141, Article 109758
6. Soumier A, Sirigu A (2020) Oxytocin as a potential defense against Covid-19?
7. Wang P et al (2015) Oxytocin-secreting system: a major part of the neuroendocrine center regulating immunologic activity. J Neuroimmunol 289:152–161
8. Gao Y, Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science 368(6492):779–782. https://doi.org/10.1126/science.abb7498
9. Chen Y, Liu Q, Guo D, Emerging coronaviruses: Genome structure, replication, and pathogenesis. J Med Virol
10. https://medical-dictionary.thefreedictionary.com/immunity
11. Karki G (2018) Immunity and its types: innate and acquired immunity
12. Five factors that affect the immune system. Available https://www.kemin.com/na/en-us/blog/human-nutrition/five-factors-that-affect-immune-system
13. Zielen S, Duecker RP, Woelke S, Donath H, Bakhtiar S, Buecker A, Kreyenberg H, Huenecke S, Bader P, Mahlaoui N, Ehl S, El-Helou SM, Pietrucha B, Plebani A, van der Flier M, van Aerde K, Kilic SS, Reda SM, Kostyuchenko L, McDermott E, Galal N, Pignata C, Santos Pérez JL, Laws H-L, Niehues T, Kutukculer N, Seidel MG, Marques L, Ciznar P, Edgar JDM, Soler-Palacín P, von Bernuth H, Krueger, Meyts I, Baumann U, Kanariou M, Grimbacher B, Hauck F, Graf D, Gonzalez Granado LI, Prader S, Reisli I, Slatter M, Rodríguez-Gallego C, Arkwright PD, Bethune C, Deripapa E, Sharapova SO, Lehmberg K, Davies EG, Schuetz C, Kindle G, Schubert R (2021) Simple measurement of IgA predicts immunity and mortality in ataxia-telangiectasia. J Clin Immunol, pp 1–15
14. Javaid M, Khan IH (2021) Internet of Things (IoT) enabled healthcare helps to take the challenges of COVID-19 pandemic. J Oral Biol Craniofacial Res 11(2):209–214
15. Gluskin E (2020) Two simple comments regarding coronavirus. Med Hypothesis 141:109788
16. Caminati M et al (2020) BCG vaccination and COVID-19: Much ado about nothing? Med Hypothesis 144. https://doi.org/10.1016/j.mehy.2020.110109
17. Schield M (2020) Coronavirus covid-19: an analysis slides 0308
18. Ding W et al (2020) Corporate immunity to the Covid 19 pandemic. National Bureau of Economic Research
19. Simonsen L, Chowell G, Andreasen V, Gaffey R, Barry J, Olson D, Viboud C (2018) A review of the 1918 herald pandemic wave: importance for contemporary pandemic response strategies. Ann Epidemiol 28:281–288
20. Gorvett Z (2020) The people with hidden immunity against Covid-19. Available https://www.bbc.com/future/article/20200716-the-people-with-hidden-protection-from-covid-19
21. Miller A et al, Correlation between universal BCG vaccination policy and reduced morbidity and mortality for COVID-19: an epidemiological study. https://doi.org/10.1101/2020.03.24.20042937
22. https://data.cdc.gov/NCHS/Provisional-COVID-19-Death-Counts-by-Sex-Age-and-S/9bhg-hcku/data
23. European Centre for Disease Prevention and Control (2020) Sonraí faoi chróinvíreas COVID-19. Publications Office. https://doi.org/10.2906/101099100099/1

24. Hatchett RJ, Mecher CE, Lipsitch M (2007) Public health interventions and epidemic intensity during the 1918 influenza pandemic. Proc Natl Acad Sci U S A 104:7582–7587. https://doi.org/10.1073/pnas.0610941104
25. Newman K, Study: high blood pressure, obesity are most common comorbidities in COVID-19 patients. Available https://www.usnews.com/news/healthiest-communities/articles/2020-04-22/obesity-hypertension-most-common-comorbidities-for-coronavirus-patients
26. Goodwin M (2020) What is a coronavirus? Available https://www.healthline.com/health/coronavirus-types
27. Jain A, Coronavirus—what is it? How does it spread? What are the symptoms and cure? Available https://www.indiatimes.com/trending/human-interest/coronavirus-symptoms-cure-medicines-sars-505173.html
28. Simpson RJ (2020) Exercise, immunity and the COVID-19 pandemic. Available https://www.acsm.org/blog-detail/acsm-blog/2020/03/30/exercise-immunity-covid-19-pandemic
29. https://www.worldometers.info/coronavirus
30. Wiegner L, Hange D, Björkelund C, Ahlborg G (2015) Prevalence of perceived stress and associations to symptoms of exhaustion, depression and anxiety in a working age population seeking primary care—an observational study. BMC Fam Pract 16:38
31. Nadeem S (2020) Coronavirus COVID-19: available free literature provided by various companies, journals and organizations around the world. J Ong Chem Res 5(1):7–13. Document ID: 2020JOCR37. https://doi.org/10.5281/zenodo.3722904
32. COVID-19 and the correctional environment: the American prison as a focal point for public health
33. Wang B, Li R, Lu Z, Huang Y (2020) Does comorbidity increase the risk of patients with COVID-19: evidence from meta-analysis. Aging 12(7)
34. Mondal SK, Bhattacharya S, Sen DJ (2020) Coronavirus: COVID-19 is now a pandemic. Acta Sci Pharmacol (ASPC) 1(5):1–2
35. COVID-19 immunity: 15% immune instead of expected 40% in Sweden; what immunological dark matter means and more. Available https://www.firstpost.com/health/covid-19-immunity-15-immune-instead-of-expected-40-in-sweden-what-immunological-dark-matter-means-and-more-8701191.html
36. Deaths, https://www.worldometers.info/coronavirus/coronavirus-age-sex-demographics/
37. Wells DW (2020) The good (but not great) news about T Cells and Herd immunity. Available https://nymag.com/intelligencer/2020/08/reasons-for-covid-19-optimism-on-t-cells-and-herd-immunity.html
38. Kissler S et al (2020) Social distancing strategies for curbing the COVID-19 epidemic. https://doi.org/10.1101/2020.03.22.20041079.t
39. Ferguson NM et al (2020) Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand
40. Andrews JL, Foulkes L, Blakemore SJ, Peer influence in adolescence: public health implications for COVID-19. Trends Cogn Sci CelPress Rev 24(5):585–587
41. Shi Y, Wang Y, Shao C et al (2020) COVID-19 infection: the perspectives on immune responses. Cell Death Differ 27:1451–1454. https://doi.org/10.1038/s41418-020-0530-3
42. Wolfel R et al (2019) Virological assessment of hospitalized cases of coronavirus disease. MedRxiv, https://doi.org/10.1101/2020.03.05.20030502
43. Nadeem S, Coronavirus COVID-19: available free literature provided by various companies, journals and organizations around the world
44. Moccia F, Gerbino A, Lionetti V et al (2020) COVID-19-associated cardiovascular morbidity in older adults: a position paper from the Italian society of cardiovascular researches. GeroScience 42:1021–1049. https://doi.org/10.1007/s11357-020-00198-w

A Comprehensive Solution for Handling Security Issues with Seaport IoT Systems

Thirumurugan Shanmugam, Mohamed Abdul Karim Sadiq, and Sudha Senthilkumar

Abstract In the current epoch, the Internet of Things (IoT) can be reflected as an important technological revolution related to evolution of smart cities, smart homes, IoT-controlled factories, and IoT for logistics in seaports implementations. With the existence of smart sensing systems in seaports becoming a reality today, different sectors in seaports are working toward a programmed mode. Some of the eye-opening projects related to smart seaports in the IoT era can be found all over the globe. In many of these new architecture implementations, even though the rapid development of IoT enables us to inspire new research works, the challenges in IoT also grow equally in terms of security. Encryption plays a key role in safeguarding IoT hardware and the data from various sensors. The proposed work focuses on reality study on various security issues emerging in the usage of IoT in seaports and suggestions for handling the security issues. Highly secure algorithms need developed in IoT encryption level standards are discussed here. Conclusions regarding the extension of future research prospects in the IoT systems and high-level security in seaports are guided in the final segment of the paper.

Keywords IoT applications · Seaport security · IoT logistics · Encryption standards

T. Shanmugam (✉) · M. A. K. Sadiq · S. Senthilkumar
Department of Information Technology, University of Technology and Applied Sciences, Suhar, Sultanate of Oman
e-mail: thirumurugan.soh@cas.edu.om

M. A. K. Sadiq
e-mail: abdulkarim.soh@cas.edu.om

S. Senthilkumar
e-mail: sudha.soh@cas.edu.om

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_52

1 Introduction

The growth of export and import services is probably glowing as a booming industry because of the rapid progression in demand for logistics services, such as maritime transport and seaport services. Such logistics require that the performance of shipping and seaports is measured to replicate the level of overall economic development. In fact, the impact of changes in financial activity and especially those of stock trade is typically intensified in logistics services.

Recent developments necessitate improved protection of the port infrastructure. Software systems that support critical operations are becoming increasingly prone to cyber-attacks. Apart from protecting sensitive data, the overall control on the operational aspects is essential. Innovative physical and cybersecurity mechanisms must be created to prevent and respond to all potential threats in these areas. Port authorities play a vital role in the international trade and economic development. The need for protecting transport infrastructures is valued by most countries.

After the crisis in 2009, competing the position, many container carriage companies were incorporated vertically with port terminals, neighborhood logistic operators, and shipping agencies. Along with the need to remain competitive over other modes of transportation, volumes of cargo also demand faster execution of regulations and better synchronization of operations. Digitalization is reflected to be important in streamlining administrative processes, allowing efficient management of cargo flows through exchange of information on cargo, organization, and equipment. Increased pressure on environmental capitals has already insisted on expressive action to subsidize the maintenance of shipping environment. Though organizational security and infrastructural flexibility represent balancing elements of a comprehensive risk management strategy, the two perceptions are different, but the vital situation is to consider both. Security is the capability to prevent adverse situations in the industry, while flexibility is the ability to reduce the magnitude and forceful distraction on current works.

The IEEE standardized the things that relate to Internet of Thing (IoT) as a network or its collection combined to form a layer where the physical devices and real-time world share the data [1]. In recent years, IoT is created as a dominant platform with increased number of sensors coupled with the developments in sensor industry. The sensors and IoT platforms create a critical data transfer adapting the changing environment transfer the data efficiently to the other devices connected with same IoT platform. The growth of high-level sensor devices has spiked, and it is estimated to maintain tremendous growth rates in the future. Most of the future occupied government projects are based mainly on the IoT platform.

The usage of sensors is dominating in all kinds of industry sectors such as power grid, designing smart buildings to monitor the thermal, and energy consumption or dissipation in walls, even to transfer the data among smart buildings. The role of sensor works out well in most revenue producing industry. Sensors are useful in the seaports and marine services. Automated or smart seaports connect all the sensor devices and control devices in a common platform with encrypted passcode with secure systems. Seaport IoT system consists of devices such as wireless devices, data transfer modems, to communicate the port information securely into the common IoT platform connected to the Internet. Seaport authorities have the required security-related data which will prioritize the needs of the customers. It is the job of a port authority to provide the required offers and services to the port customers in a fast and precise way through IoT devices. Our research focuses on the following proposed novel contributions.

- To design a smart seaport, the entire seaport IoT architecture is designed using latest technology sensors such as ultrasound sensor, eddy current sensor, infrared sensor, intrusion detection sensors, motion detection sensors, vibration alerting sensors, and weight detection load cells.
- Those sensors are connected into a controlling platform for level conversions and normalization of data. These data from IoT sensors accurately control the devices in a faster way.
- Obstacle sensing is used for automatic-guided vehicles in seaports which suggest the proper positioning of the vessel when it travels using high-beam lasers and encoders.
- Another seaport job is moving the huge containers from one place to another through heavy cranes.
- Automated IoT-controlled system provides the fast handling of such loads and displace the same to the right place in ships using machine learning algorithms.
- Smart parking system, billing and code marking for the parcels, detection of unknown vehicles, pedestrian tracking in seaports and anonymous roaming in seaports and ships, security alerts and smart energy management systems are some attractions of the proposed work.

2 Review of Literature

In 2014, Skarmeta et al. [2] delivered a hybrid algorithm that is integration of AES and ECC algorithms. Encryption using AES and ECC is enhanced for security purpose. New challenges in IoT systems were also discussed. In 2013 [3], cipher text-based system is developed to provide optimum security structure in IoT networks. Handling the cipher-level inputs is the beginning for IoT security system development. This worked out deeply for increasing the security of encryption keys which is a regular IP-controlled security system despite developing an optimized model. The novel system is developed almost every year to deal with security challenges in IoT. Hence, the level of consistent growth in securing the data and connectivity keeps on strengthening every year. Authors in [3] and [2] created a way for us to start investigating and understanding the security traits of the IoT system and provide good idea on where the research needs to be started.

In 2015, Xin [4] created a clear view in IoT cybersecurity systems. The users of Internet are growing rapidly every year in an increasing scale. Millions of new users enter the Internet world. The usage almost provides us the comfort, and users benefit in a high level; most of the unsecured platforms act as information gathering networks. The Web sites grab your information such as name, register numbers, email ids, passwords, personal information, and passwords. People use to provide the information in the Internet platform easily and faster the data will spread over the unsecure networks. The cybersecurity system creates a strong wall between the hackers and the users to protect the data to be hacked by the unknown, but still, the equivalent increase in the development of hackers and their ability to hack the information beyond the cybersecurity also grows. Xin [4] discussed the challenges in cybernetics and how to utilize the IoT devices properly from the hackers were discussed in his research work which also motivated us to move the study toward the cybersecurity as concern.

In 2018, authors [5] investigated a research study on efficient usages of IoT in creating smart logistics. The increased level of imports and exports in India is obviously act as one of the economic strengths to the country. Implementing the smart system everywhere becomes unavoidable and provides the customers the required comfort on transactions, etc. Logistics is one of the rapid growing industry and busier day and night for the import and export activities. Storing the large amount of user data in the cloud also sometimes becomes overwhelming. Statistically logistics industry generates more data day by day, and obviously, keeping the data safely in the secured cloud is a major challenge. The authors [5] discussed about the smart logistics system using smart connectivity. Every package is handled wirelessly in a hustle-free manner.

In 2012, Zeng [6] states that IoT-based system and mainly discusses about the impact of 4th-generation communication system in the Internet users. The research work [6] focuses on 4G and the things of Internet, creation of impact and pitfalls in data security, etc. The users got rapidly increased when the 4th-generation network comes on the way because of the fast data rate, attractive benefits, and hustle-free services everywhere. Peoples started using the network more and more for online purchase, banking, advertising, cloud storage, video processing, and medical image processing, uploading the information faster into the cloud. The growth in 3rd-generation increased like a sudden burst to 4th generation because of the enormous abilities present in it.

In 2011, authors [7] investigated more about hardware involved in things of Internet. The success of IoT is based on the parameters like the security level, processing speed, and data transfer rate, using the appropriate embedded hardware which is never influenced by malfunctions. Authors [7] focus the research work on investigating the hardware devices used for IoT, troubleshooting capabilities of hardware when the hackers malfunction the security system also the reliability of the hardware system for the prolonged working condition beyond the environmental changes and disasters. Hardware suitable for IoT will be based on proper device selection and weightage of the processor working on it. IoT transmitters and receivers are well encrypted through qualified programmers and mutual dependency to provide higher security. In his research, he found out that intelligent data transfer is possible when the processor self-repair its issues coming over the cloud and avoiding stack to be overflowing. The management of processing the data in a pipelined manner or parallel manner matters the speed.

In the year 2017, authors [1] conducted a study on cybersecurity risk levels in maritime industry. As the study states that maritime industry is heavily exposed to cybersecurity attacks, and risk management is not properly managed there. There are so many different maritime vehicles like boats and ships which are less equipped and handled by different operating background. The communication with the ship, environmental changes, and disaster updating each and every information traveled through various sensor nodes is connected with unique protocols. The study focuses on security challenges faced by the industry and concluded with a security encryption scheme with improved optimization in the year 2017.

In 2018, authors [8] concluded the research work on fusion ports. Combining the operations of one or more seaport security systems through a common cloud to hybrid the structure for high-level security is analyzed by the authors [8]. Cybersecurity is a vital apprehension for the functioning of a modern economy. There is a crucial need for progress accepting of microeconomic mechanism in the cybersecurity industry and for consistent data upon which policy design can be based: this with reference to transportation sector, where the goal of secure data communication can only be achieved by a strong and continuous collaboration between public and private sector. Every day, cybersecurity software demonstrates unsuccessful in stopping attacks. Malware, malicious users, and embattled attacks create enormous risk. For industry, government and consumers to appreciate cybersecurity success, they must begin with a groundwork built on trusted hardware completing.

3 Design Methodologies Adopted

The nation's growth depends on the reputed industries providing their contribution and support to build the overall economic growth of the country. The exports and imports of various goods keep on increasing year on year, which created the seaport services to the peak level of engagement. An IoT system helps the seaports to communicate the information rapidly and precisely to avoid manual time lags.

The growth of IoT-enhanced systems creates easy access to things which connect the people and devices through a customized platform connected with Internet. Network providers act as an agent for connecting the things with the devices. The enormous growth in IoT-enabled systems also created challenges in seaport sectors to provide required security of user data that are stored in the cloud servers.

The growing challenges in security traits, hacking of information, illegal usage of information enable the researcher to create more secure algorithms on encryptions and decryptions. These created a demand for more hybrid algorithms on data encryptions beyond the standard AES and DES schemes. The security is enhanced by including random chippers and pattern-based chippers too. Focusing on tuning, the encryption method provides the IoT system coming up with new modified versions. Security based on user typed text, face encryption technology, bio-inspired encryption and tracking system, unique multi-modal biometrics are also being used.

3.1 The Data Encryption Standard (DES)

The data encryption method is a standardized model of creating mutually dependent keys encrypted in a fixed model of steps. The number of iterations of key generation is also fixed. Both the sides, sender or the user and the receiver or the server system keep the unique private key. The further process is known as a symmetric key algorithm.

3.2 Advanced Encryption Standard (AES)

AES uses a single encryption key of varying lengths. The AES algorithm concentrates on a single block of data and re-encrypts it 10–14 times, depending on the key length. When using an Internet-connected medical device, AES meets US Government requirements for HIPAA data protection. AES also meets FINRA standards for protecting financial records. AES is an efficient and elegant algorithm whose strength resides in its key length options. The longer the key length, the more exponentially difficult it is to break the encryption.

When comparing both the algorithms DES and AES, it is evident that DES is less secure than AES. In fact, DES encryption is a result of a 30-year-old effort by the US Government to offer cryptographic security for all government communications. The goal was to achieve both cryptographic security and standardization well known to everyone. DES is the main impact for cryptography but has since been split by researchers.

3.3 Triple Data Encryption Standard (3DES)

An interesting algorithm created after the analysis and pitfalls of AES and DES algorithm is a type of computerized cryptography where each block of data receives three passes of secure secret data. Enormous security comes from the larger key length. Triple DES was also replaced by NIST, which adopted the well declared algorithm called AES. Triple DES is now considered obsolete but is still used by some IoT products worldwide because of its compatibility and flexibility. Triple DES algorithm secures the data well and protects against brute force attacks. Brute force attacks use formulated tools to find out various combinations of secured data until the hacker cracks the key.

3.4 RSA Encryption

As the study continues, another algorithm comes on the way in which initials RSA come from the last names of three founders of RSA Data Security (Rivets, Shamir, and Adelman). RSA encryption provides a public-key encryption technology licensed by RSA Data Security, who also sells its accompanying development kits.

RSA encryption scheme is intelligent and allows users to send encrypted data without having to share the code with the recipient. It is a public-key encryption, and the public key can be shared openly. However, the data can only be decrypted by another private key known by the end users. Each RSA user has the common public key, but only, designated recipients are privy to the private key.

3.5 Twofish Encryption Algorithm

Twofish is another well-defined block cipher algorithm proposed by Counterpane Labs over 20 years ago as a replacement for the AES. Because of the enormous benefits and security, Twofish algorithm was a finalist for selection as the new NIST advanced encryption standard. Twofish applies a block ciphering system based on a single key of any length up to 256 bits. This encryption standard is more efficient on computers with lower magnitude capacity of system processors and IoT device smart cards. Twofish appears in many of the free encryption software products like Vera Crypt.

3.6 Encrypting IoT Systems

The study enforced to declare a new systematic method which involves better security challenges in Internet of Things system applicable in all industries. Development of machine learning algorithms tangles the sparks on adding more accuracy and hacking free security protocol in IoT modules. Machine learning algorithms are more uniquely modeled, which analyze the inputs in the form of image, text, or an audio clip, or the input will be a video sequence with different objects. The system model is divided into training and testing. This is the method of considering numerous statistical possibilities of a certain pattern of inputs with the comparative analysis of patterns, features, style, color, and other objectives of the group which is in the global datasets.

The analysis of machine learning in a particular model considers globally connected data, collected records of environment, registered information of customers in large scale, etc. The analysis is obviously complex and seems to be a complex model in understanding. The source of information collected through seaport IoT modules is from mobile devices, cloud information, etc. The seaport

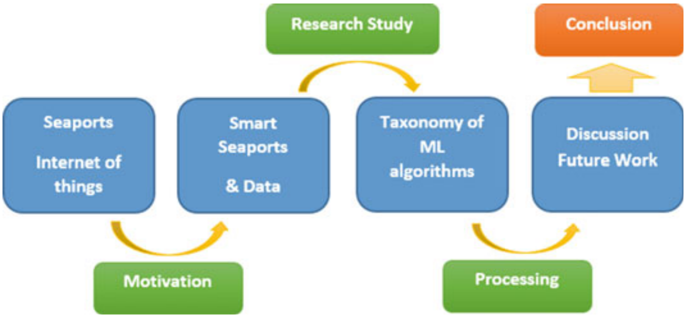


Fig. 1 Organization of the study

customers login daily to the authorized network provided by seaports and share the information on common cloud.

The combination of ML-enhanced algorithms and hybrid encryption systems is used to make the encryption more precise and overcome the security pitfalls in seaport networks. Another advantage persists in combining the ML algorithms to be used in IoT modules is clarifying the unpredictable hacking challenges, information lost in network, pretend the unconditional behavior of IoT modules during processing.

The deep understanding of ML protocols is required to overcome the similarity issues in duplicating the encrypted data. Reliable system is demandable in seaport networks to avoid the overwhelming problems of handling big data vector and fast loading data during registrations. Encryption is doubled when algorithmically over covered system presented in the IoT modules of seaports. The usage of cameras in seaports is helpful to analyze the activities of the humans and loading/unloading activities of the containers to find out the deviations quicker; henceforth, the ML algorithms play numerous roles in seaport IoT systems and provided double-encryption security.

4 Tabulations

See Table 1.

5 Findings and Discussions

After the study of various algorithms related to IoT data encryption and selection of data through ML process such as algorithm A, B, C, D, and E which area applied to avoid the anomalies, it is evident that usage of ML in encryption for data security plays an efficient role. ML algorithms are strategically evident in the case of generating

Table 1 Comparison of various encryption algorithm performance on processing time

Research author(s)	Encryption method	Processing time in ms/MB transmission
Pavithra G—2018 [8]	AES	5.6
Amirhossein Safi—2017 [9]	DES	5.8
Kedar Deshpande 2018 [10]	Triple DES	6.3
Vidhya Vijayan, Eldo P Elias 2019 [11]	RSA	6.1
Thorat CG, Inamdar VS 2018 [12]	TWOFISH	5.6

modified model which can double encrypt the data to provide security and focused on creating such ML-combined algorithm to overcome the unpredictable behaviors of overloaded data. The study hopefully paves a way to extend the research on covering the formulated encryption algorithms with the highly secured ML prototypes to form a double-encrypted scheme for IoT-connected modules in seaport systems. The major challenge is unpredictable victims on processed data; hence, the middle-level communication module needs higher security schemes than that of existing IoT implementations.

6 Conclusion

The Internet of Things is growing with increasing pattern every year. It has spread from the power grid to smart refrigerators at home, at industrial levels by monitoring production line efficiency creating various smart systems. The devices communicate with Internet will be encrypted because of the personal and business intelligence data they transmit. IoT users are aware, and security managers focus on creating nonvulnerable algorithms to safeguard the user info residing in the large cloud. The protocols are framed in such a way that creates necessary security to confidential information through levels of encryption. The study conducted here analyzes various encryption algorithms such as AES, DES, Twofish, RSA, and Triple DES. The best protection available right now is encryption. Algorithms and elegant mathematics-based machine learning approaches are evident right now to provide required security. As far as so many discussions on algorithms mentioned in Refs. [13–19], the study is positively concluding that creating hybrid encryption schemes with the association of ML algorithms provide more security in IoT systems at the seaport.

7 Future Scope

In future, deep learning-based approaches will be implemented to improve the security of data on the cloud. Testing of real-world data can be carried out for the ports.

Acknowledgements This research is funded by Ministry of Higher Education, Research, and Innovation under the Block Funding Program—Research Grant (RG) 2020.

References

1. Bou Harb E, Kaisar EI, Austin M (2017) On the impact of empirical attack models targeting marine transportation. In: 5th IEEE international conference on models and technologies for intelligent transportation systems. https://doi.org/10.1007/978-1-4614-1860-3_2
2. Skarmeta AF, Hernández-Ramos JL, Moreno MV (2014) A decentralized approach for security and privacy challenges in the Internet of Things. In: IEEE conference "World Forum on Internet of Things, pp 67–72
3. Hong N, Xuefeng Z (2013) A security framework for internet of things based on SM2 cipher algorithm. In: International conference on computational and information sciences, Shiyang, pp 13–16
4. Xin M (2015) A mixed encryption algorithm used in internet of things security transmission system. In: International conference on CEDC and knowledge discovery, Xi'an, pp 62–65
5. Arumugam SS, Umashankar V, Narendra NC (2018) IOT enabled smart logistics using smart contracts. In: 8th International conference on logistics, informatics and service sciences
6. Zeng L (2012) A security framework for internet of things based on 4G communication. In: 2nd International conference on computer science and network technology, Changchun, pp 1715–1718
7. Babar S, Stango A, Prasad N, Sen J, Prasad R (2011) Proposed embedded security framework for Internet of Things (IoT). In: 2nd International conference on wireless VITAE, Chennai, pp 1–5
8. Pavitra G, Deshpande SL (2018) Security for VITA in IOT using AES algorithm. Int J Sci Technol Eng 5(1):76–77
9. Safi A (2017) Improving the security of internet of things using encryption algorithms. Int J Comput Inf Eng 11(5):558–561
10. Deshpande K, Singh P (2018) Performance evaluation of cryptographic ciphers on IoT devices. Glob J Eng Sci Res [ICRTCET'18], pp 418–427
11. Vijayan V, Elias EP (2019) Hybrid method for securing data in IoT cloud. Int J Innov Technol Explor Eng (IJITEE) 8(5):447–450. ISSN: 2278-3075
12. Thorat CG, Inamdar VS (2018) Implementation of new hybrid lightweight cryptosystem. Appl Comput Inf 16(1/2):195–206
13. Hosseinpour F, Abu Bakar K, Hardoroudi AH, Kazazi N (2010), "Survey on artificial immune system as a bio-inspired technique for anomaly based intrusion detection systems. In: International conference on intelligent networking and collaborative systems
14. Michel CD, Thomas PF, Tucci AE (2017) Cyber risks in the marine transportation system. FAC/Documents/USCG_Paper_MTS_CyberRisks.pdf?ver=201-07-19-070403-473
15. Pishva D (2016) IoT: their conveniences, security challenges and possible solutions. ICACT Trans Adv Commun Technol 5(2):797–808

16. Schauer S, Stamer M, Bosse C, Pavlidis M, Mouratidis H (2017) An adaptive supply chain cyber risk management methodology. In: Hamburg international conference on logistics. https://doi.org/10.15480/882.1491
17. Mc Nicholas M (2016) Maritime security: an introduction. Butterworth-Heinemann
18. Xu Z, He J, Chen Z (2012) Design and actualization of IoT-based intelligent logistics system. In: IEEE international conference on industrial engineering and engineering management, Hong Kong, pp 2245–2248
19. Awan IA, Shiraz M, Hashmi MU, Shaheen Q, Akhtar R, Ditta A (2020) Secure framework enhancing AES algorithm in cloud computing. In: Security and communication networks, vol 2020, Article ID 8863345, pp 1-–16

Dynamic Load Balancing in Cloud Network Through Sunflower Optimization Algorithm and Sine–Cosine Algorithm

U. K. Jena, Pradipta Kumar Das, M. R. Kabat, and Sanjay Kumar Kuanar

Abstract The objective of the paper is to scheduling of independent task dynamically among the virtual machines (VMs) in the cloud network on the share resources. Scheduling of task and allocating of the resources from the data center have been performed through the several meta-heuristic algorithms and achieved encouraging results. However, their performance evaluation is far based on the ideal state and needs more improvement. Load balancing in necessary when some VMs are executing more number of task and task need to wait in queue for processing at the same time other VMs are free and not allocated any task or less number of tasks for execution. The problem under consideration is proposed sunflower optimization algorithm with sine–cosine algorithm (SFOA-SCA) for improving the performance of load balancing in cloud network. The experimental result illustrates that the projected procedure is outstripping its opponent in the manner of throughput time, waiting time, response time, execution time, and utilization energy during load balance of task in cloud network.

Keywords Virtual machines (VMs) · SFOA-SCA · Waiting time · Response time · Execution time

U. K. Jena (✉) · M. R. Kabat
Department of Computer Science and Engineering, VSSUT, Burla, Odisha, India
e-mail: uttamkumarjenasir@gmail.com

M. R. Kabat
e-mail: kabatmanas@gmail.com

P. Kumar Das
Department of Information Technology, VSSUT, Burla, Odisha, India
e-mail: daspradipta78@gmail.com

S. K. Kuanar
Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India
e-mail: sanjay.kuanar@giet.edu

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_53

1 Introduction

Load balancing is a challenging task in the cloud computing which allows the user to access service such as sharing of resource, database, software, and server on demand basis [1]. In the cloud computing environment, user submitted request to task manager for scheduling of task. The task manager provides a request to the scheduler for scheduling the task and assigning to VMs for execution. Many scheduling algorithm has been proposed by the researchers to resolve the load balancing issue in the cloud network. Heuristic approach has been proposed for task scheduling and resource allocation in cloud computing environment [1]. The author proposed bandwidth aware division scheduling and BAR optimization for resource allocation. Dragonfly optimization and constraint measure have been exercised for load balancing in cloud computing [2]. Hybridization of firefly and improved multi-objective particle swarm optimization (PSO) have been illustrated for energy-efficient load balancing [3]. The author proposed a hybridization of meta-heuristic algorithm by modifying PSO with Q-learning for dynamic load balancing in cloud computing [4]. The aim of hybridization was to improve the throughput, energy utilization, and optimizing of waiting time. An efficient version of binary PSO algorithm with low time complexity and low cost for scheduling has been proposed to resolve the load balancing and resource allocation in cloud environment [5]. Cloud task scheduling has been carried out through the enhanced SFOA [6]. Whale optimization algorithm has been exercised for cloud task scheduling with multi-objective optimization model [7]. The aim was to improve the performance of a cloud system with a specified resource. Hybridization of electro search with genetic algorithm has been proposed to resolve the task scheduling in cloud computing [8]. Modified hennery gas solubility optimization and oppositional based learning has been illustrated to resolve the task scheduling in cloud network [9].

The objective of the paper is to generate an effective and efficient load balancing in cloud network by distributing the task among all VMs equally. Load balancing is carried out by designing the fitness function for VMs with considering four different constraints. The overall fitness function evaluates before assigning any task to VMs and generates the optimal VM through the integrated SFOA and SCA. SFOA algorithm has been improved and integrated in the SCA to generate the best position of VMs for assigning task to overcome the overloaded and under loaded VMs. Integration SFOA algorithm into SCA is to enhance the convergence speed and to provide the better optimal solution in the search space and also avoid to trap at local optima.

The rest of the paper is illustrating as follows: The design of the fitness function for load balancing is elaborated in Sect. 2. Sunflower algorithm is described in Sect. 3. Section 4 described the sine–cosine algorithm and also algorithm for load balancing in cloud computing network. Experiment setup and result analysis has been illustrated in Sect. 5. Conclusion of the work is described in Sect. 6.

2 Design of Fitness Function for Load Balancing

The main emphasize of the proposed technique for load balancing is to distribute the task among all VMs in equal manner as to reduce the power consumption, migration cost, and utilization of memory during the load balancing. The fitness function is designed to evaluate the capability of VMs for executing the task with a minimum makespan. Consider M number of PMs serves the resource pool in the cloud network which is owned K number of resources and each PM consists of N number of VMs. Let P_i and VM_{ij} signify the ith $(1 \leq i \leq M)$ PM in the cloud and jth $(0 \leq j \leq N)$ VM in ith PM, respectively. C_{ik} signifies the total capacity of k-type the resources owned by the ith PM. D-dimensional vector is used to represent the request for VM for executing the task and sharing of the resources. The fitness function is designed based on the makespan, energy utilization, and power consumption and migration cost and memory usages. The first fitness function is designed based on the minimum makespan. The execution time of the jth VM is T_j, and it is calculated based on the decision variable U_{ij}.

$$U_{ij} = \begin{cases} 1, & \text{if } T_i \text{ is assigned to VM}_j \\ 0 & \text{if } T_i \text{ is not assigned to VM}_j \end{cases} \tag{1}$$

$$T_j = \sum_{i=1}^{n} U_{ij} \times \mathrm{TC}_{ij} \tag{2}$$

where TC_{ij}

where TC_{ij} is the ith $(1 \leq i \leq n)$ task completion time in jth $(1 \leq j \leq N)$ VM and TC_{ij} is calculated as follows:

$$\mathrm{TC}_{ij} = \frac{L_i}{\mathrm{PS}_j} \tag{3}$$

where L_i is the length of the ith task and length of the task is defined in terms of number of instructions (Millions of instruction) and PS_j is the processing time of jth VM in the cloud. The makespan (MS) is the maximum value of the execution time of all virtual machines and mathematically defined as follows:

$$F_1 = \mathrm{MS} = \mathrm{Max}(T_j), \quad 1 \leq j \leq m \tag{4}$$

The second fitness function is designed based on the power consumption during the load balance and it is determined by taking the absolute Euclidean distance all dynamic PMs simultaneously. The smaller Euclidean distance indicates the better load balancing and PM is turn off mode when there is no task executing in the related

PM. The efficiency of power (E_f) of each active PM is calculated as follows:

$$E_f^i = \sqrt{\sum_{j=1}^{k} (\mathrm{RU}_j - \mathrm{RUbest}_j)^2} \tag{5}$$

where RU_i signifies the jth resource utilization and resource may be memory and CPU, and RUbest_i represents the best utilization of jth resource for power efficiency in each PMs. The power efficiency of all PMs at time is calculated as follows:

$$P_t = \sum_{i=1}^{M} E_f^i \tag{6}$$

The total power efficiency of the system within T is expressed as follows:

$$F_2 = \sum_{t=0}^{T} P_t \tag{7}$$

The third fitness function is designed based on the migration cost of VM, and it is expanded when the quantity of movements increments. The better load balancing provides the least movements of task from the loaded VMs to unloaded VMs. The migration cost is calculated as follows:

$$F_3 = \frac{1}{N}\left(\frac{\sum_{i=1}^{N} \text{Number of migration in } \mathrm{VM}_i}{N}\right) \tag{8}$$

The fourth fitness function is designed based on the memory usages, and the usages of the resources in VMs are memory and CPU. The computation of storage usages in whole cloud arrangement is evaluated as follows:

$$F_4 = \frac{1}{M * N}\left[\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{1}{2}\left(\frac{\text{CPU Utilized}_{ij}}{\mathrm{CPU}_{ij}} + \frac{\text{memory utilized}_{ij}}{\text{memory}_{ij}}\right)\right] \tag{9}$$

The overall fitness function is designed for load balancing in cloud network by taking the weighted sum of the individual fitness function designed based on the constraint evaluations. The fitness function for load balancing in cloud network is expressed in Eq. 10.

$$F = \beta_1 F_1 + \beta_2 F_2 + \beta_3 F_3 + \beta_4 F_4 \tag{10}$$

3 Sunflower Optimization Algorithm

The sunflower optimization algorithm (SFOA) mimics the movements of the sunflower to fascinate the solar radiation [10]. It generates the optimal solution based on two principles such as pollination and movement. In the pollination phase, sunflower liaises to harvest the pollen gamete. In movement phase, the sunflower proceeds toward the best sunflower in a random manner which is considered as sun here. Every optimization algorithm needs the proper balance between exploration and exploitation to generate the better global optimal solution on the search space. To explore this issue, basic SFOA has been improved by introducing a novel pollination operator that maintains the good balance between exploration and exploitation capability. The procedure of improvement is explained bellow:

Initialization of Population: Population (P) is initialized for N sunflowers, and each sunflower consists of M numbers. The population is initialized through as follows:

$P = \{X_1, X_2, X_3, \ldots, X_N\}$ and each sunflower $X_i = \left\{x_{i,1}, x_{i,2}, \ldots, x_{i,M}\right\}$ and $x_{i,j}$ is signified in the following manner:

$$x_{i,j} = x_{\min} + \beta_{ij}(x_{\max} - x_{\min}) \tag{11}$$

where β_{ij} signifies the random number in the range of [0,1], $x_{\max}$ and $x_{\min}$ are the upper and lower bound of the sunflower $x_{i,j}$ correspondingly and its value is set as M and 1 correspondingly.

Pollination: In the pollination process, the classical SFOA select pr $\times N$ sunflowers from the population, where pr the pollination is rate, and then, update each nominated sunflower as follows:

$$X_i(t+1) = \lambda_i\big(X_i(t) - X_j(t)\big) + X_j(t) \tag{12}$$

where $X_i(t)$ and $X_j(t)$ are the position of the sunflower i and j at iteration t, respectively. p is the pollination rate, and λ_i is the random number within the range [0,1]. The pollination process is modified in the improved SFOA as follows:

$$X_i(t+1) = \alpha(t)Z + (1 - \alpha(t))Y \tag{13}$$

$$Z = X_j(t) + \text{rand}\,(X_i(t) - X_j(t)) \tag{14}$$

$$Y = X_i(t) + \delta(X_{\text{best}}(t) - X_i(t)) \tag{15}$$

$$X_{\text{best}}(t) = \min(f(X_i)) \;\; \text{where,} \;\; 1 \le i \le N \tag{16}$$

$$\alpha(t) = \alpha_{\max} - \frac{t}{T}(\alpha_{\max} - \alpha_{\min}) \tag{17}$$

$$\delta = 2 \times \left(1 - \frac{t}{T}\right) \tag{18}$$

where $\alpha(t)$ is the control parameter which control the local and global pollination ration and its value lies in 0 and 1, δ is the scaling factor which regulates the amplitude of the searching direction $(X_{\text{best}}(t) - X_i(t))$, $X_{\text{best}}(t)$ is the best position of the sunflower found so far, and rand is the random number with in [−1,1]. T is the maximum iteration and t is the current iteration. The value of $\alpha_{\max} = 0.7$ and $\alpha_{\min} = 0.3$. The man motive of introducing novel pollination process is to improve the exploration and exploitation capability simultaneously. In improve SFOA, at the start of the iteration, it performs more on the global search and proceeds toward the local search with the iterations. The local pollination process is carried out in Eq. 13 through the term Z which explores the exploitation process in the search algorithm and the term Y introduced in Eq. 13 enhances the exploration process by global pollination.

Position Update: The SOFA is very much sensitive on the pollination rate. The high sensitivity does not guarantee to provide the better sunflower. In this improvement, the pollination rate has been modified as follow:

$$\text{pr} = 0.5 \times \left(1 - \frac{t}{T}\right) \tag{19}$$

The position is updated based on the selection of pr $\times$ N numbers of sunflowers and updated their position as follows:

$$X_i(t+1) = X_i(t) + \lambda \frac{X_{\text{best}}(t) - X_i(t)}{\|X_{\text{best}}(t) - X_i(t)\|} \tag{20}$$

The position update is carried out to change the direction of the sunflower toward the sun $X_{\text{best}}(t)$.

4 Sine and Cosine Algorithm (SCA)

It is a population-based meta-heuristic algorithm to generate the optimal solution in the search space by generating the initial random solutions and allow them to proceed toward the most promising solution in the search space through the mathematical function sine and cosine [11]. Nevertheless, the population-based algorithm concentrated on two phases such as exploration and exploitation to obtain the optimal solution. The position is updated through the classical SCA as in the following manner.

$$X_i(t+1) = \begin{cases} X_i(t) + p_1 \times \sin(p_2) \times \left| p_3 P_i^g - X_i(t) \right| , & p_4 \prec 0.5 \\ X_i(t) + p_1 \times \cos(p_2) \times \left| p_3 P_i^g - X_i(t) \right| & p_4 \geq 0.5 \end{cases} \tag{21}$$

where p_4 is the control parameter for selecting sine or cosine function to update the position of the particles in the search space and the value of the control variable lies within [0,1]. The parameters p_1 and p_2 select the direction of movement and amplitude of the movement in the search space, respectively. The parameter p_3 is random weights which determine the effect of the goal as the iteration proceeds. p_4 is used, the greedy strategy, to control the balance between exploration and exploitation through the selection of sine or cosine function. The classical SCA agonizes from slow convergence rate, more time taken to obtain the optimal solution, trapped at local optima, and it proceed to engender the global optima based on the adaptive and randomness control of the parameters. Therefore, it is unable to generate the satisfactory solution in each time because of fine tuning the control parameters. Hence, it is necessary to modify the classical SCA through $X_{\text{best}}(t)$ which is the best position of SFOA and used the greedy strategy to update the position of the particle based on the SOFA or modified SCA to avoid the slow convergence and trapped at local optima in the search space. The modification of SCA as follows:

$$X_i(t+1) = \begin{cases} X_{\text{best}}(t) + p_1 \times \sin(p_2) \times |p_3 X_{\text{best}}(t) - X_i(t)| , & p_4 \prec 0.5 \\ X_{\text{best}}(t) + p_1 \times \cos(p_2) \times |p_3 X_{\text{best}}(t) - X_i(t)| & p_4 \geq 0.5 \end{cases} \tag{22}$$

The parameter p_3 is the random number lies in between [0,1] and p_1 and p_4 are modified as follows:

$$p_1 = \mu\left(1 - \frac{t}{T}\right) \tag{23}$$

where μ is positive constant and its value is set as 0.7.

$$p_4 = 1 - \frac{t}{T} \tag{24}$$

Algorithm 1: Pseudo code for SFOA-SCA for load balancing

Input: N is the Population; T is the maximum allowable iterations and initializes the population $X_i(1 \leq i \leq N)$ using Eq. 33. $X_{\max}$ and $X_{\min}$ are the upper bound and lower bound of the search space, respectively. Initialize the parameters of SFOA and SCA, N number of VMs, M number of Tasks.
Output: Optimal Load balancing in cloud Network
Initialize the population through Eq. 11
While($t < T$).
1.Set the best position of sunflower as X_{best}.
2.For $i = 1\ to\ N$
3. Compute the fitness value of X_i.
4.End for
5.X_{best} = minimum($f(X_i)$).
6.For i=1 to N
7. $(X_i)_{SFOA}$ is the position updated through Eq. 20.
8. $(X_i)_{SCA}$ is the position updated through Eq. 22
9. If($f\left(\left((X_i)_{SFOA}\right) \prec f\left((X_i)_{SCA}\right)\right)$ Then
10. X_i =$(X_i)_{SFOA}$
11. Else
12.X_i =$(X_i)_{SCA}$
13. for End
14.while End
15.Return global best solution

5 Experiment Setup and Result Analysis

The load balancing in cloud network has been exercised through the CloudSim 4.0 simulator. The simulation has been carried out in the Desktop machine consists of Intel Quard-core i7processor, 3.4 GHz CPU, 16 GB RAM, 1 TB HDD, and Window 10 platform. The environment setup for carrying out the experiment has been presented in Table 1.

The performance of the algorithm verified through the result discussion and result of the load balancing has been exercised in terms of response time for task, makespan, number of task migrated, resource utilization, throughput, and degree of imbalance. The experiment has been conducted in real and simulation environment. The outcomes of the experiment have been presented in Table 2. The outcomes of the result reveal that the proposed algorithm outperforms over its counterpart algorithm in different metrics. The performance of the proposed algorithm has been evaluated in the terms of resource utilization during load balancing and its outcome has been illustrated in Fig. 1. The outcomes of the Fig. 1 show that SFOA-SCA exploited the maximum number of resources during the load balance as compared to its competitive algorithms. The energy utilization during the load balance through the SFOA-SCA with its counterpart algorithm has been presented in Fig. 2. The result of Fig. 2

Table 1 Environment set for load balancing

Type	Number	Parameters	Value
Data center	2	Arch	×86
		OS	Linux
		VM Monitor	Xen
VM	10–100	Processor speed	9726MIPS
		Memory	0.5 GB
		Bandwidth	1 GB/s
		Image size	10 GB
		Number of PEs	1
		VM Manager	Xen
		VM policy	Time shared
Host	20	Processing element power	10,000,00MIPS
		Storage	4.0 TB
		RAM	16.0 GB
		Bandwidth	15 GB/s
		Cores	6
		Storage capacity	2 TB
		Total number of tasks	1000
		Length of each task	1 MB–2 GB
		Total file size as task	50 GB

illustrates that the proposed algorithm uses minimum amount of energy as compare to other algorithms. Makespan has been evaluated through SFOA-SCA, SFOA, and SCA during the load balancing in the cloud network and it is illustrated through Fig. 3. Figure 3 reveals that SFOA-SCA takes less time to execute the assigned tasks in VMs as compared to other algorithms. Number of task migrated during the load balance has been evaluated for maintaining the equilibrium of load balance in VMs, and outcome of the result has been presented in Fig. 4.

Figure 4 shows that the number of task migrated among the VMs is least as compared to its counterpart algorithms. Hence, the proposed algorithm outperformed as compared to its counterpart algorithms. The response time of task has been performed through the proposed algorithm along with its counterpart algorithms. The response time of the algorithms for load balancing has been illustrated through Fig. 5. Figure 5 shows that the performance of the algorithm in respect to the response time is minimum as compared to other algorithms. Throughput of the load balancing has been evaluated through the proposed algorithm, and it is presented through Fig. 6.

Figure 6 illustrated that the proposed algorithm provides the better throughput as compared to other algorithms in its counterpart. Therefore, the proposed algorithm

Table 2 Comparison result of real environment and simulation environment

In Real Platform Environment							In Simulation Environment			
VM	TASK	Algorithm	Makespan (in ms)	No. of Task migrated	Throughput time (in ms)	Energy utilization (in KJ)	Makespan (in ms)	No. of task migrated	Throughput time (in ms)	Energy utilization (in KJ)
16	100	SFOA-SCA	2113.73	17	4693.84	82.75	2325.10	18	5163.22	91.02
	200		2237.11	22	4718.72	87.39	2460.82	24	5190.59	96.12
	300		2289.83	25	4749.28	89.96	2518.81	27	5224.20	98.95
	400		2318.27	29	4771.51	92.11	2550.09	31	5248.66	101.32
	500		2348.43	34	4803.87	94.55	2583.27	37	5284.25	104.00
	100	SFOA	2178.73	27	4976.12	87.40	2396.60	29	5573.25	97.88
	200		2237.37	32	5048.38	89.74	2461.10	35	5654.18	100.50
	300		2275.69	43	5076.85	92.56	2503.25	47	5686.07	103.66
	400		2315.17	52	5121.07	94.03	2546.68	57	5735.59	105.31
	500		2335.68	65	5156.15	96.18	2569.24	71	5774.88	107.72
	100	SCA	2979.14	35	5158.82	90.95	3277.05	38	5881.05	103.68
	200		3045.26	42	5191.48	93.22	3349.78	46	5918.28	106.27

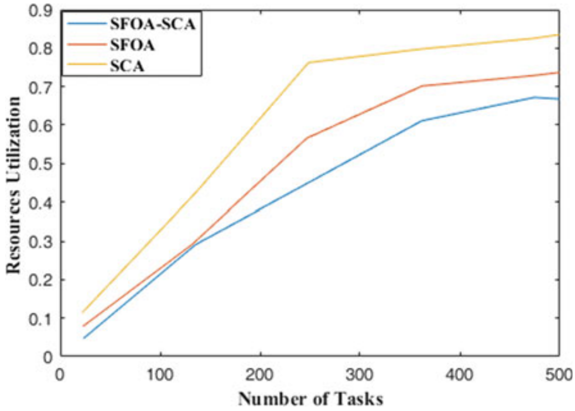


Fig. 1 Number of Tasks with Resource utilization for fixed number of VMs

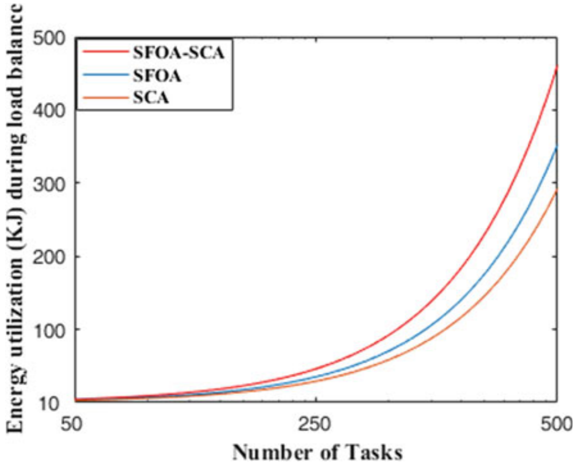


Fig. 2 Number of Task with Energy utilization during load balancing for fixed number of VMs

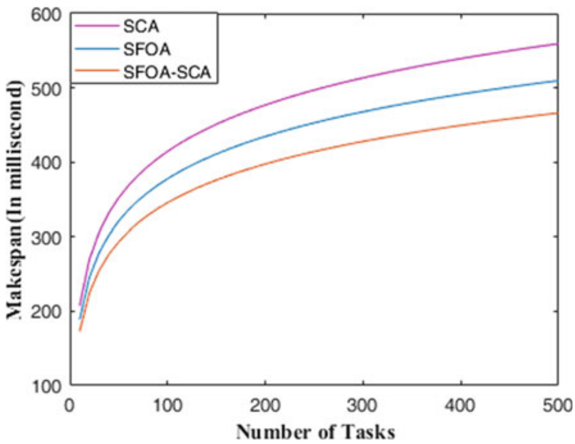


Fig. 3 Number of Tasks executed with Makespan for SFOA-SCA, SFOA and SCA

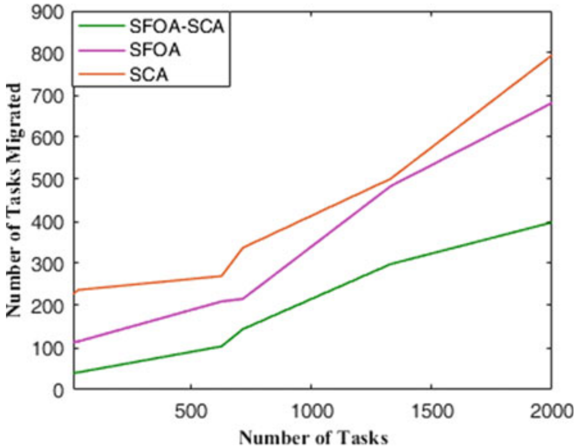


Fig. 4 Number of task versus Number of task migrated during load balance using SFOA-SCA, SFOA and SCA

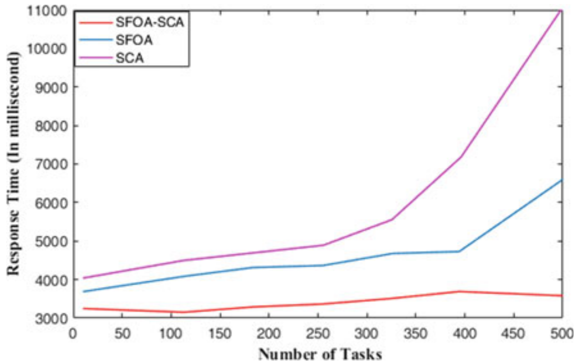


Fig. 5 Number of Task with response time of VMs using SFOA-SCA, SFOA and SCA

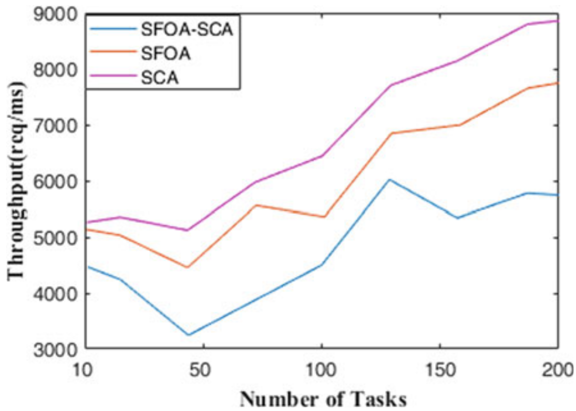


Fig. 6 Number of Task with throughput using SFOA-SCA, SFOA and SCA

SFOA-SCA generates the better results in respect to the response time, throughput, energy utilization, task migration, makespan, and resource utilization as compared to other algorithms in its counterpart.

6 Conclusions

Load balancing of independent task in cloud computing presented in this paper has been carried out using SFOA-SCA. Fitness function has been designed for assigning the task into VMs by considering different constraints. Before assigning the task into VMs, it evaluates the fitness value to check whether the tasks are overloaded or under loaded in VMs. The overloaded task will have migrated to under loaded VMs to maintain the balance of load in VMs. The performance of the proposed algorithm has evaluated through different metrics and also compared it with counterpart algorithms. The conclusion revels from the analysis of the result that the proposed algorithm is more effective and efficient for dynamic load balancing as compared to its counterpart algorithms.

References

1. Gawali MB, Shinde SK (2018) Task scheduling and resource allocation incloud computing using a heuristic approach. J Cloud Comput Adv Syst Appl 7(4):1–16
2. Polepally V, Shahu Chatrapati K (2019) Dragonfly optimization and constraint measure-based load balancing in cloud computing. Cluster Comput 22(1):1099–1111
3. Devaraj AF, Elhoseny M, Dhanasekaran S, Lydia EL, Shankar K (2020) Hybridization of firefly and improved multi-objective particle swarm optimization algorithm for energy efficient load balancing in cloud computing environments. J Parallel Distrib Comput 142:36–45
4. Jena UK, Das PK, Kabat MR (2020) Hybridization of meta-heuristic algorithm for load balancing in cloud computing environment. J King Saud Univ Comput Inf Sci
5. Mapetu JP, Buanga ZC, Kong L (2019) Low-time complexity and low-cost binary particle swarm optimization algorithm for task scheduling and load balancing in cloud computing. Appl Intell 49(9):3308–3330
6. Emami H (2021) Cloud task scheduling using enhanced sunflower optimization algorithm. ICT Express
7. Chen X, Cheng L, Liu C, Liu Q, Liu J, Mao Y, Murphy J (2020) A woa-based optimization approach for task scheduling in cloud computing systems. IEEE Syst J 14(3):3117–3128
8. Velliangiri S, Karthikeyan P, Arul Xavier VM, Baswaraj D (2021) Hybrid electro search with genetic algorithm for task scheduling in cloud computing. Ain Shams Eng J 12(1):631–639
9. Abd Elaziz M, Attiya I (2021) An improved Henry gas solubility optimization algorithm for task scheduling in cloud computing. Artif Intell Rev 54(5):3599–3637
10. Ferreira Gomes G, da Cunha SS, Carlos Ancelotti A (2019) A sunflower optimization (SFO) algorithm applied to damage identification on laminated composite plates. Eng Comput 35(2):619–626
11. Das PK (2020) Hybridization of Kidney-Inspired and sine–cosine algorithm for multi-robot path planning. Arab J Sci Eng 45(4):2883–2900

Machine Learning Approach to Exploratory Data Analysis on Global Terrorism

Debabrata Dansana, Subhashree Sahoo, Faizan Ashraf, Vivek Kumar Prasad, and Kalyan Das

Abstract This article provides an overview of the benefits and the limitations of using exploratory data analysis for research questions surrounding explosion incidents. Most importantly, it provides a starting place for those interested in learning about this type of analysis and considerations that users should be aware of. Those unfamiliar with exploratory data analysis may find the article helpful in understanding what will be covered during their dissertation process. Terrorism is a term we use to refer to violence and intimidation in the pursuit of political aims. Terrorism has been around for about two centuries, but it did not gain much traction until the 1970s. Terrorism is defined as violent acts against innocent people committed to making an impact on society or government. The underlying causes of terrorism are usually motivated by injustice, revenge, oppression, retaliation, conflict, and hostility borne out of inequality among countries. Furthermore, it is associated with religious extremism and usually targets civilian populations to cause death or destruction. As a result of these patterns in manifestation and potential consequences worldwide, terrorism has become a pressing global issue that requires attention at multilateral levels. In this article, we have tried to cover all the aspects of terrorist data analysis using various machine learning algorithms like support vector machine and linear regression. After analyzing the above datasets by the said model, we have figured out which region is mainly affected by attacks and where the cases of terrorists have been seen frequently. Here, we have tried to explain what it means to do this kind of

D. Dansana (✉)
Rajendra University, Balangir, India
e-mail: debabratadansana07@gmail.com

S. Sahoo
Centurion University of Technology and Management, Bhubaneswar, India

F. Ashraf
GIET University, Gunupur, India

V. Kumar Prasad
Nirma University, Ahmedabad, India
e-mail: vivek.prasad@nirmauni.ac.in

K. Das
Sambalpur University Institute of Information Technology, Sambalpur, India

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_54

research and provide some considerations for those who want to use this approach for their work.

Keywords Terrorism · Support vector machine · Linear regression · Prediction · Data analysis

1 Introduction

Terrorism is defined as "the systematic use or threat of violence by an individual or group to create fear (i.e., terror) to achieve a political, religious, or ideological goal." According to the "United States Code Title 18 Chapter 33, Section 233-2a," this is the definition. The United Nations defines Terrorism as "premeditated, politically motivated violence perpetrated against non-combatant targets by sub-national groups or clandestine agents." Analyzing international crime data can lead us to understand why this type of behavior happens. Terrorism is a term we use to refer to violence and intimidation in the pursuit of political aims. Terrorism has been around for about two centuries, but it did not gain much traction until the 1970s. Terrorism is defined as violent acts against innocent people committed to making an impact on society or government. The underlying causes of terrorism are usually motivated by injustice, revenge, oppression, retaliation, conflict, and hostility borne out of inequality among countries. Furthermore, it is associated with religious extremism and usually targets civilian populations to cause death or destruction. As a result of these patterns in manifestation and potential consequences worldwide, terrorism has become a pressing global issue requiring multilateral attention. Terrorism is defined as the deliberate and systematic murder, mayhem, and threatening of innocent people to spread fear among adversaries or non-combatants. It is not a new phenomenon: Terrorism had existed since 1835 when western European states sent mercenaries into Algeria to help French landowners restore order after a long and bloody rebellion by Algerian Muslims who wanted independence from France; however, it did not come fully into its own until later in history [1].

The country that suffers from terrorism is less likely to have positive views toward people who are Muslim. For the study, researchers surveyed nearly 500 Americans. The team found that one-third of people with direct exposure to terrorism-related acts and injuries report feeling more negative toward Muslims. More exposure to terrorist violence was associated with more negative views. However, they found that direct exposure was not significantly associated with more negative sentiments toward the Muslim religion. On the other hand, people with indirect exposure were substantially more likely to report holding unfavorable views of Islam, which was not directly associated with their exposure to terrorism. The authors attributed this disparity to the fact that direct exposure to terrorism was associated with increased negative views about the United States, and indirect exposure was not. Lead author and political scientist Michael A. Salter Assistant Professor of politics and international studies at George Washington University, said, "our findings show the effects of terrorism on the mental health of Americans." The team also found that the more a person witnessed a mass shooting, the more likely they held opposing views toward Muslims. Salter said, "to people who say the United States is safer now, we can say the mental health of the American people is the highest, and it has been since immediately following 9/11." "It might be difficult to separate one's reaction to terrorism from the reaction to the policies of the current political leadership," Salter added. As for possible reasons for people's negative views, the researchers pointed to the media's tendency to highlight terrorism in certain areas, such as the media's coverage of the Boston Marathon bombings. Nowadays, terrorism is a global issue. Every country is making its plans for dealing with it. United nations have also taken significant steps to deal with it. The world is facing a new kind of disaster. It seems that terrorism will be a common problem in the world. The countries which have suffered terrorism are Pakistan, Iraq, Algeria, Iraq, Afghanistan, Yemen, Nigeria, Sudan, Israel, India, Iran, Kenya, Libya, Sudan, Saudi Arabia, Palestine, Syria, Morocco, Indonesia, Egypt, Yemen, Kuwait, Indonesia, Libya, Iran, Nigeria, Uganda, and the list of the above is too long and too difficult to enumerate. When the number of terrorism acts has increased, the security apparatus, intelligence, and police have faced a big challenge and taken a few steps to control it. This problem is not only one of the countries; it has now become a global problem. To know about the terrorist acts are essential. It is to prevent such actions. You cannot kill all the terrorist acts, but you should know about them and deal with them properly to prevent such an attack. There are various ways to prevent terrorist acts. In this article, we will tell you about some of the ways that may help in dealing with such actions. To keep an eye on any suspicious person [2].

2 Literature Review

Green Arther Sandag [3] study is aimed to conduct an exploratory data analysis and predict terrorist activity in Indonesia using K-nearest neighbor and k-fold cross-validation. In this research, data selection, data cleaning, and data reduction were carried out, and the feature selection process aimed to find out the most influential data attributes. Mullins [4] examines the culture and dynamics of terrorist organizations, the operation of such organizations in the United States, and terrorist tactics and strategies. Other chapters consider weapons used by terrorists, counter-terrorist options, victims of terrorism, hostage situations, and the future of terrorism [4]. In his paper, Singer [5] illustrates how to use interpretable classification algorithms to identify subgroups of terrorist incidents that share common characteristics and result in mass fatalities. ARIMA [6] and SEIR [7] technique are used to analyze the COVID-19 predictions and analysis in addition to that machine learning approach is used in [8] to forecast social media. In addition to this methodological contribution, from a practical perspective, exploring the factors identified in the patterns can lead to prevention strategies, such as alteration of the physical or systemic environment. We present three examples of terror attacks described by a design that resulted in a high probability of mass casualties. In contrast, attacks that differ in just one of these characteristics resulted in far fewer casualties. We propose an exploration of the differentiating characteristic to reduce the probability of mass-fatality terrorist incidents. Pravallika [5] has described and predict the region and country of a terrorist attack using machine learning approaches. The work has been carried out upon the Global Terrorism Database (GTD), and an open database containing a list of terrorist activities from 1970 to 2017.

The Global Terrorism Database (GTD) documents more than 200,000 international and domestic terrorist attack worldwide since 1970. The GTD is a compilation of terrorism information collected and organized from the world's most extensive collection of terrorism-related sources. Including the Federal Bureau of Investigation (FBI), the National Counterterrorism Center (NCTC), the British Home Office, the Israeli Ministry of Foreign Affairs, the U.S. State Department, and other agencies. A summary of the data collection process is provided below: The GTD contains data on all types of terrorist attacks: violent, nonviolent, and mass atrocities, including hostage-taking, attempted, and thwarted Terrorism. The GTD is organized by year, country, region, and type of terrorism involved. With details on various dimensions of each episode, the GTD familiarizes analysts, policymakers, scholars, and journalists with patterns of terrorism that can inform response and prevention strategies. Data collected and contained include the following:

Country: The full name of each country.

Region: A country's subdivision into geopolitical regions.

Year: Data are organized by year and include information on the country, region, type of attack, and the date and time of each attack.

2.1 Primary and Secondary Data Sources

The primary data source for the GTD is the FBI's National Database (NDB) of Domestic and International Terrorism. The NDB contains detailed information on all known and suspected terrorist incidents in the United States and abroad, regardless of the motive or number of victims. This database includes more than 30,000 cases or approximately 60,000 cases from 2003 to 2009. The GTD also uses data compiled and coded by the FBI's National Criminal Information Center (NCIC). The NCIC contains numerous other reports on terrorism, including the NCIC database of arrests and court convictions of terrorists. In addition, the NCIC contains information on terrorism arrests and court convictions of foreign citizens and nationals. For information on more than 30,000 people arrested in the United States, including both domestic and foreign nationals, see the NCIC Terrorists and Terrorist Organizations database. The GTD also includes additional data compiled and coded by the Joint Terrorism Task Forces (JTTF). The JTTFs are federal, state, and local multi-jurisdictional collaborative investigations. These law enforcement agencies identify terrorism in progress and work together to disrupt and prosecute terrorists. The GTD contains reports compiled and coded by JTTFs. See Terrorism in the United States [1, 9].

2.2 Data Quality

To assure the quality of the GTD, the FBI and the various JTTFs conduct ongoing quality control procedures. Data that do not meet FBI and JTTF standards are removed from the NDB. In addition, all the information included in the GTD is coded by FBI and JTTF personnel. During coding, any inconsistencies or omissions are flagged and subsequently investigated and corrected. If the records contain evidence of significant errors, the documents are referred to the FBI's Assistant Director for Intelligence (ADBI).

In addition, the FBI also conducts internal audits of GTD records for compliance with FBI procedures and standards. These audits are performed by designated groups of FBI personnel who are familiar with the subject matter and can detect unusual or erroneous information. Audits identify unnecessary deletions of records and records that appear to be erroneously coded or entered, duplicate, and outdated or stale records. If necessary, the FBI corrects the records. However, if the changes are minor and not a significant source of error, the difference is allowed to stand as a correction to the original document. Recent terrorist attacks that devastated many forms of lives in the past. Some of the significant incidents took many lives and all. This creates a global issue as well as it needs to stop. The leading causes of terrorism are poverty, the conflict between states, and a lack of tolerance. The first and most important method to stop terrorist acts is getting everyone involved and working together to prevent them. The second most important method is to avoid wars and stop the fights between countries that can cause terrorism. The third and most important method is to treat terrorists as they are. They are not all good, and they are not all bad; they are human. If a country can stop it, it should, and the world should also stop if it cannot. There is no such thing. "War" in a broad sense means "conflict over scarce resources," and "war" in a more specific sense means "two or more armed groups engaging in conflict." There is no conflict; that is a subset of war. The only wars that can be considered wars are officially recognized by the U.N. and have international sanctions. The U.N. can be broken down into several different categories, depending on how they are formally and legally defined in international law. These include formal treaties, usually signed by several states, and binding international law. These are relatively easy to define since they are generally characterized by a set of binding rules, which can be enforced. The rules can be codified in law or custom, but either way; these rules define the legal framework in which the conflict takes place non-international law. These include unilateral acts that are not formally binding, like treaties or declarations. These are harder to define since they are more difficult to enforce and thus not everyone considers them binding. One can make a good argument that everything is non-international law since any act by any government on another government or its citizens is non-international. While definitions could be used to distinguish between actions, wars, and conflicts, none of them help identify the borders between them. Some regions are greatly served as the hot zone of terrorism. From these regions, significant people have joined the terrorist groups and participated in several terrorist attacks. In this part, some of the major terrorist organizations are discussed in detail [4, 10].

3 Problem Statement

Here, we have tried to identify the major causes of terrorism and hot zones by analyzing the global terrorism dataset using machine learning technique.

3.1 Proposed Work

Here are the various factors that we have considered for analyzing our dataset and explain the results of these processes by a suitable graph.

Analyzing dataset while considering the timeframe

Given that dates sort the data, attacks on U.S. citizens seem rare in each date range. But the terrorist act against the citizens of the U.S. has been increasingly in the following year after this rare date range. By finding the date of the start of the increase, the factors in increasing terrorist acts can be easily identified by taking into account the changes and developments in the country after this date. For example, we find the increase of terrorist acts in 2010 and the beginning of this increase in 2009. If we were to look at the factors of increasing terrorism from 2009 to 2010, we could see that most of the factors that caused increasing terrorism were already in place before 2009. But after the year 2010, the effect of this cause is seen in rising terrorism. Using a linear regression model, we find the factors that were the leading cause of terrorism in the years from 2005 to 2010: In the equation, $y1$, $y2$, $y3$, and $y4$ are the values of variables: date, year, month, and day, respectively. The model explains the year-by-year increases of the value of the *y*-axis up to the value of 3.16 from the base year, 2005, which illustrates an 80% increase in the value of the *y*-axis (the year-by-year data are available from 2005 to 2010). When looking at the graph in Fig. 1, it is possible to see that the increase in the value of the *y*-axis started at the beginning of 2009 (2009.16), which means the value of the *y*-axis is increasing by 0.16 each year. The increase of terrorism acts from 2005 to 2010 can be explained by the linear regression model. (*y*-axis shows the value of terrorism effects, the *x*-axis shows the year, and $y1$–$y4$ are values of the model). Based on this model, we will attempt to predict the potential impact of the event that will be considered an event that could increase the values of $y1$, $y2$, $y3$, and $y4$, and the difference of $y1$–$y4$ can measure the impact. In the model, the value of $y1$ (the variable that shows the effect of factor 1) is a product of factors 1–4, so the value of $y1$ could be expected to be the product of the factors. When factoring one increases by 0.1, factor 2 will be 1.0 minus 0.1, and factors 3 and 4 will be unchanged. Based on this model, the year of 2009.16 is the base year (a year without the increase of the factor), and then, the year

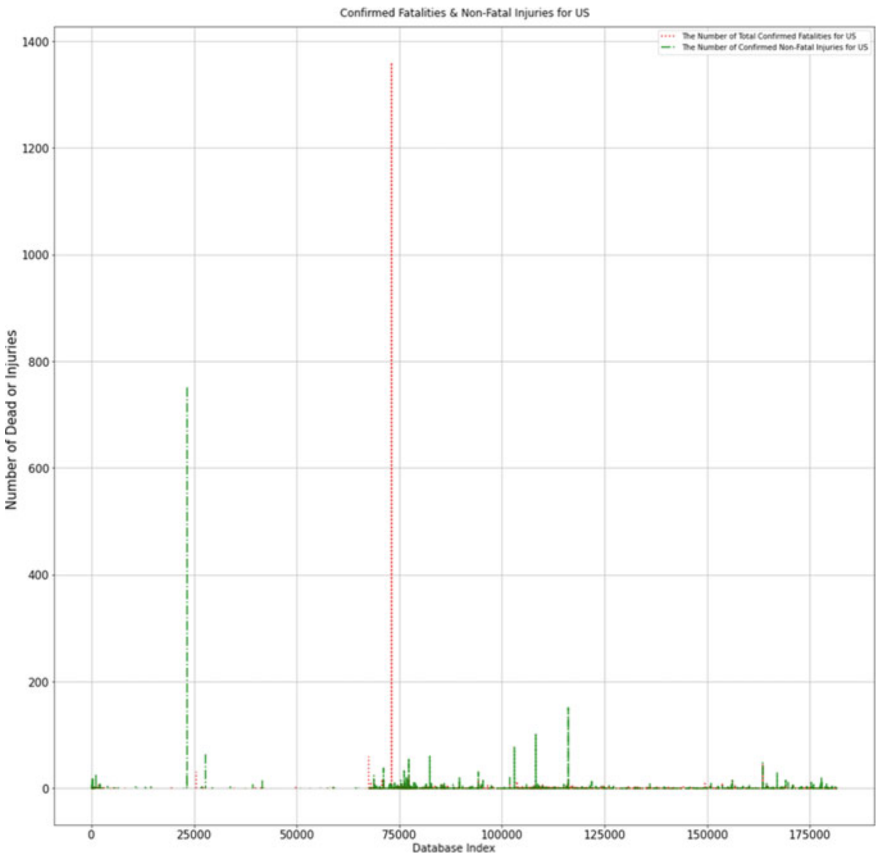


Fig. 1 Mortality report for terrorist attack on U.S. in a certain timeframe

of 2009.25 is the predicted year (with the rise of the factor), and this is supposed to have a different value than $y4$. Then, the differences of $y1$–$y4$ (the predicted value of $y1$) are shown as x-axis on the graph in Fig. 1.

A simple linear regression function can be written as

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \ldots, n$$

We can obtain n equations for n examples:

$$y_1 = \beta_0 + \beta_1 x_1 + \epsilon_1 y_2 = \beta_0 + \beta_1 x_2 + \epsilon_2 \vdots y_n = \beta_0 + \beta_1 x_n + \epsilon_n$$

If we add n equations together, we get

$$\sum y_i = n\beta_0 + \beta_1 \sum x_i + \sum \epsilon_i$$

Because for linear regression, the sum of the residuals is zero. We get

$$\sum y_i = n\beta_0 + \beta_1 \sum x_i \tag{1}$$

If we use the ordinary least squares method, which aims to minimize the sum of the squared residuals. We define C to be the sum of the squared residuals:

$$C = (\beta_0 + \beta_1 X_1 - Y_1)^2 + (\beta_0 + \beta_1 X_2 - Y_2)^2 + \cdots + (\beta_0 + \beta_1 X_n - Y_n)^2$$

3.2 Support Vector Machine

Support vector machine" (SVM) is a supervised machine learning algorithm that can be used for both classification and regression challenges. However, it is mainly used in classification problems. In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is several features you have), with the value of each element being the value of a particular coordinate. Then, we perform classification by finding the hyperplane that differentiates the two classes very well and is close to the points in the positive and negative classes. SVM is an efficient algorithm for linear problems in which all the data are linearly separable, and the input dimensions are less than the number of the features. As the dimension of the input data increases, the probability of the existence of the non-linear relationship between the elements and the classes becomes high. Since this problem can be very high-dimensional, it is better to try non-linear classifiers such as the K-nearest neighbor algorithm, decision tree, linear-discriminant analysis, Gaussian network, and other non-linear algorithms. To use a support vector machine (SVM) represented in Table 1, you should implement the training algorithm of the support vector machine, which is the kernel trick. Here is a simple algorithm that performs the SVM training:

Table 1 SVM algorithm

Input: A given training data points D = (X_1, Y_1) …, (X_m,Y_m) with m instances in d-dimensional space Output: A vector fv in the feature space (d + k) dimensional space
Step 1. Normalize each training instance as (x_1-mean) …, (x_m-mean) and (y_1-mean) …,(y_m-mean) Step 2. Compute the kernel trick f: D - > Rn Step 3. Create a function g: Rn - > R1 Step 4. Initialize all the parameters of SVM as: ξ, ξ^2, β = 1 Step 5. Compute u_i = g(x_i) + ξβ(f(x_i) - y_i) for i = 1…, m Step 6. If u_i is negative for some i, remove that instance and update the parameters (ξ, ξ^2, β) Step 7. Train the SVM Step 8. Generate the SVM classification rule as: f(x_i) = + 1 if u_i > 0; f(x_i) = −1 if u_i < 0

Form of equation defining the decision surface separating the classes is a hyperplane of the form:

wTx + b = 0.

– w is a weight vector.

– x is input vector.

– b is bias.

Allows us to write.

$wTx + b \geq 0$ *for di* $= +1.$

$wTx + b < 0$ *for di* $= -1.$

The easiest way is to make a recursive function. Something like this:

```
def bin (x, n):
   if n == 0:
      return []
   else if x == −1:
      return [−1] * (n − 1)
   else:
return [bin (x − 1, n − 1) + [−−] + bin (x + 1, n − 1)]
```

4 Results and Discussion

4.1 Analyzing Dataset While Considering the Mortality Rate

In most acts of terrorism, the mortality rate and injuries were low, but a small number of actions led to too many deaths and injuries. Here, we represent the data using the scatter plot; wherein *X*-axis, we have taken the mortality rate. In *Y*-axis, we have taken the injuries that happened during terrorism. This figure shows the number of mortality and injuries that have increased over the years. Figure 2 shows the graph of mortality versus injuries in particular years. We found that the number of deaths had risen when the casualties were more than the deaths due to a specific cause. We have discussed these trends in various sections with a focus on a particular reason.

The trends are compared between two periods before 2004 and after 2004. Before the year 2004, period consists of the events that took place before 2004, then after the year 2004, period consists of the events that took place after 2004. The mortality due to terrorism rose from 0.00% to 3.3% from 1998 to 2007. The injuries also increased from 1.50% to 17.5% during the same period. However, these injuries had risen from the year 2001 onwards. Before the year 2001, the rate of injuries was 1.5%, and in the years 2001–2003, it was 1.71–2.11%. We could observe a gradual increase in the damages from 2001 to 2007. From 1998 to 2007, a total of 1732 terrorist acts occurred; 817 of these were committed in the year 2001. The highest number of terrorism-related deaths was reported in the year 2001 as compared to the following years. The rate of terrorism-related mortality was found to be 3.35% in the year

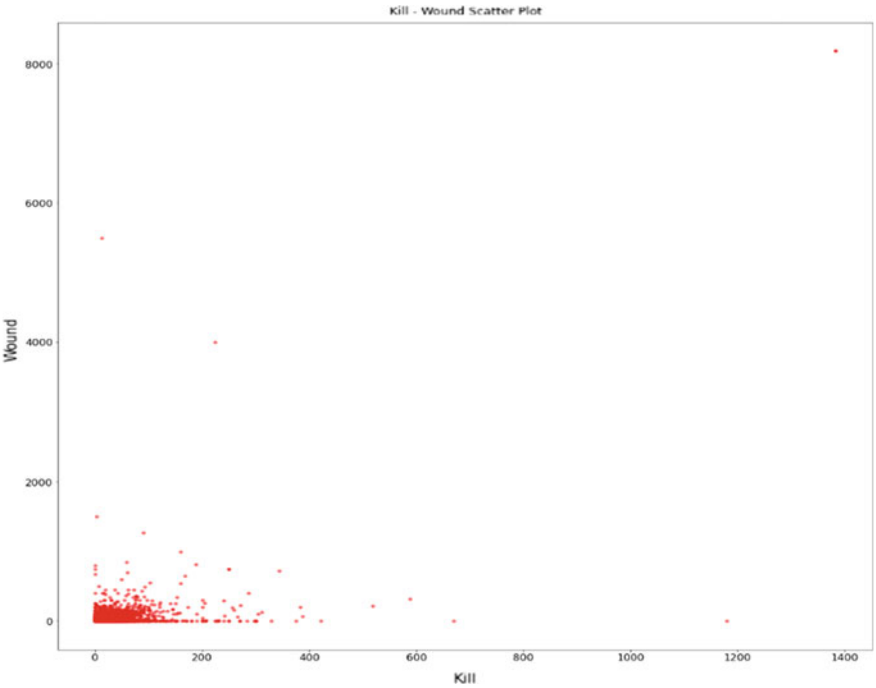


Fig. 2 Mortality and fatality report on the terrorist attack

2001. In the years after 2001, the mortality due to terrorism increased by ten times compared to 2001.

4.2 Frequency of Terrorist Actions in Customized Region

Let us analyze the Middle East and North Africa. Figure 3 gives an overview of the terrorist actions in a particular region for a specific period. We have taken the years of attacks on the *X*-axis, and the *Y*-axis provides an overview of the frequency of attacks in that period. How these terrorist attacks have been increasing in recent years can be what we get from the diagram. There is a rise in the frequency of attacks from the years 2004–2006. We also see a rise in the total number of terrorist attacks in this particular region. This is again explained by the recent wave of terror strikes that Islamic fundamentalist terrorist groups have made. We can also analyze the geographical distribution of such terrorist attacks. The following figure also shows this. We can see that most terrorist attacks have been made by fundamentalist Muslim fundamentalist organizations, which is not surprising. But such events are not restricted to this region only. Such events are even happening in some countries of Europe. For instance, in the U.K. and France in which two bombs were made in

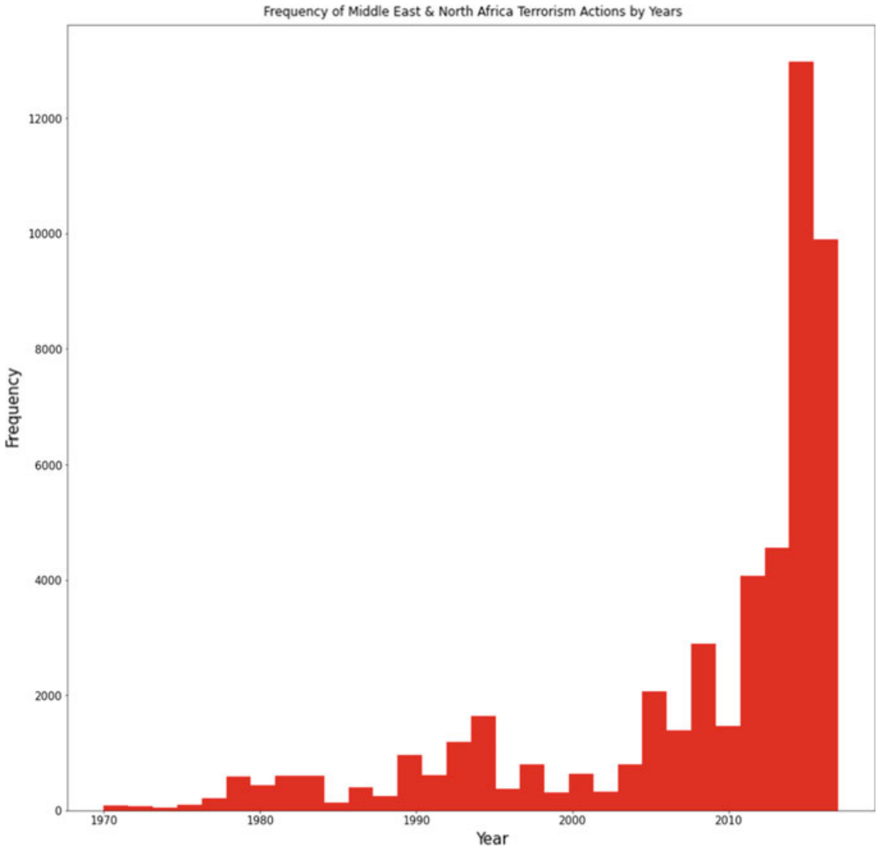


Fig. 3 Frequency of terrorist attacks on the Middle East and south region

different locations. And in Germany, there were two very similar terrorist acts. And then, there was a bombing in which a massive explosion took place in the area of Dortmund in Germany. There was also a shooting in a mall in Germany, happening in the same time frame. All these are terror attacks and are part of a more significant wave of terror. If you go to the Wikipedia link, you will get additional information regarding this particular terror wave. But the very reason for the lock is the U.S. invasion of Iraq and the consequent surge of terror.

One interesting thing is that many countries in the Muslim world and the Middle East are part of the axis of evil countries, Syria, Iran, Iraq, Algeria, Libya, Egypt, Saudi Arabia, and Kuwait. In one of the terror attacks in the U.K., a car bomb was made in Libya. Now, if you see the map below, you will know the location of these countries. Let us compare this to the countries that have invaded Iraq and have since been mentioned as the countries of the axis of evil. It is not by chance that the majority of these countries are the ones that have invaded Iraq. And for those who want to see the reason for the terror attacks. During that period, most of the war was happening.

Also, in the period in which we are in, most of the war is happening; the invasion of Iraq was on. The map below depicts the area of the charge and the areas of the terror attacks in Iraq, and the areas of the axis of evil countries. Let us compare the two maps to see which country is closest to the location of a terror attack.

As you can see from the map, the closest country to the location of the Madrid bombing and the bombing in London is Syria. As we already mentioned earlier, it is in Syria that many terrorist organizations are located. So, suppose we conclude that a lot of the terror attacks in the Middle East and Iraq are part of the axis of evil. In that case, we can also figure that most of the terror attacks in the Middle East and Iraq support the U.S. invasion and war in the Middle East and the subsequent surge in terror. Another exciting thing is that many of the terror attacks planned in these countries like Saudi Arabia, which is close to the location of a terror attack, have also been thwarted.

4.3 Terrorist Attacks of a Particular Year and Their Locations

In this Fig. 4, we analyze the terrorist attacks data over a specific year worldwide. In the *X*-axis, we have taken the years, while the *Y*-axis gives an insight into the mortality rate due to these attacks. Here, we have compared our whole data worldwide to the Middle East and North Africa terror attacks. As the pictures depicting all over the world terrorism incidents are different. But from the *Y*-axis, it appears that there is no significant variation in the incidents or mortality rate compared to other parts of the world. To see the data's interpretation and impact on the planet, we have taken the values in the *y*-axis and converted them to percent, as shown in Fig. 5. In Fig. 5, the *y*-axis values are in percentages and plotted for the years 1995–2016. As we can see, the rate is highest in the Middle East, followed by North Africa and Africa.

Further, as we investigate the graph, we can say that the rate is less in all parts. However, the rate has increased continuously over the years. It was important news that the most significant attack had occurred in Bali, where 202 people had been killed. And from this graph, we can say that if we continue to do the same way, we may not have to face this type of incident of terrorism. It may give some clue how this number of terror attacks in any other part of the world.

However, another vital point to be noted in this graph is that we can say that the death toll is highest in countries that belong to the Middle East region as these countries belong to the most religious countries in the world. In addition to that, they are also one of the worst affected when it comes to terrorist attacks. Hence, it can be noted that although they have a higher rate of mortality in the incidents, it is in terms of percent than in other parts of the world. So, we can assume that most of the terror attacks in these parts of the world do not spread terrorism. And hence, a big reason for this could be that the deaths occurring in the terror attacks are generally from their home country. It is evident from this graph that although terrorism is spreading in many parts of the world, they do not spread terrorism as such. It does not get spread to other parts of the world as we can see that it is mainly limited to specific

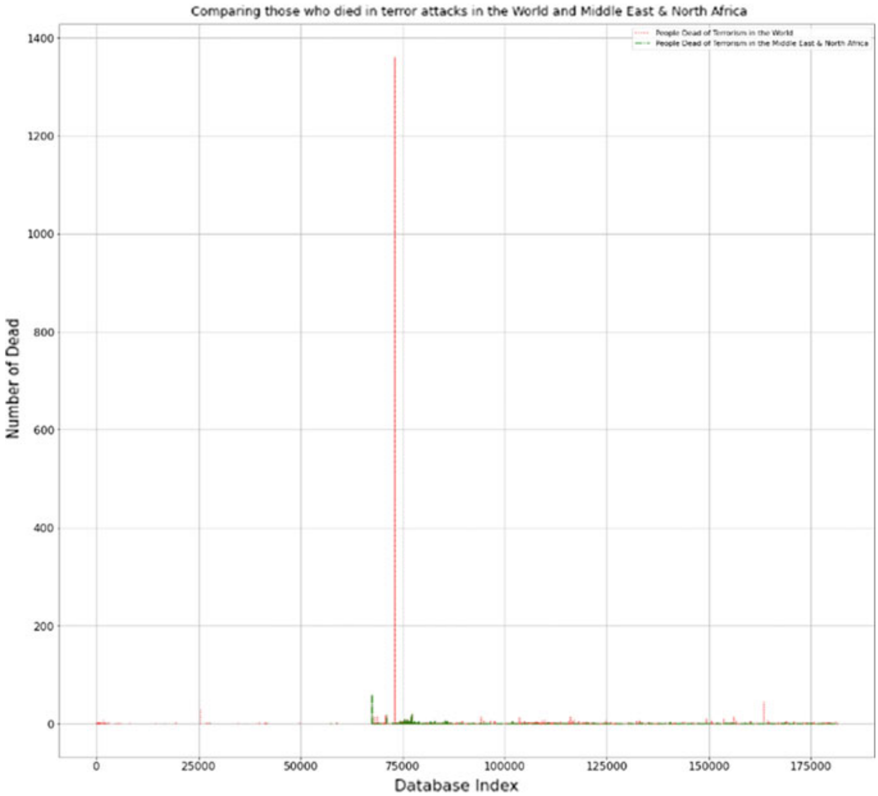


Fig. 4 Comparison of the mortality rate of the south region and Middle East with the rest of the world due to terrorist attack

countries in the Middle East and not more. As we look at this graph, we can see that in many cases, this is being done in terms of spreading terror and not for any other reasons. There are many cases when people are not afraid of terrorism, but still, they are willing to get killed because of this. This can be because of the religious reasons practiced in most of the countries in the Middle East. However, it is not the same for other parts of the world. However, there are other factors that are spreading terrorism. And we can see that the growth rate of this graph has been increasing over.

In this Fig. 5, we analyze the terrorist attacks for various reasons; data armed assault and bombing/explosion are seen to be the cause of 77% of the deaths in these attacks. Hostage, hostage tracking, and unarmed assault are the reasons, too. This rate is why these attacks are used so many times in terrorist actions. This is how dangerous weapons and explosives are to the world. In this figure, the rate of the type of terrorism is the third highest in the world. This is because most terrorist attacks use weapons to kill people. When most of the killings are made by weapons, it is a terrifying issue. In this figure, the rate of the type of terrorism is the second highest in the world. We have seen in the previous figure that the reasons for this

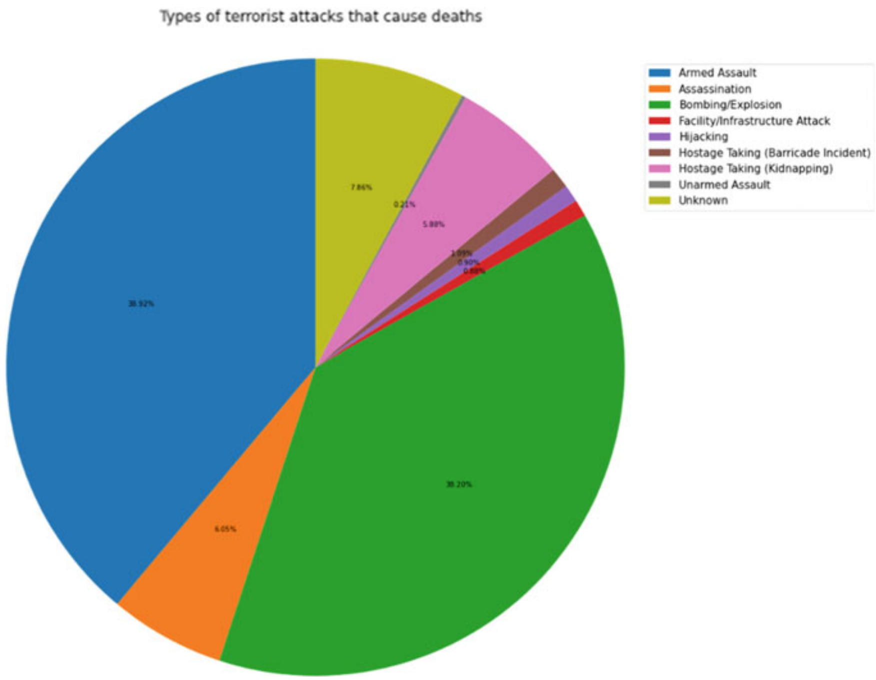


Fig. 5 Pie chart showing medium of the terrorist attacks

action are armed assault and bombing/explosion. In this figure, the type of terrorism is why civilians have died so much because these reasons caused many deaths and caused many injuries to civilians. This is like the type of terrorism we cannot ignore anymore, and it also gives us an idea of what is going on in the world now. This is like the type of terrorism we cannot ignore anymore, and it also gives us an idea of what is going on in the world now. The kind of terrorism I talk about is apparent in this figure. From this figure, we can conclude that most of the terrorist actions are done by Islamic militants. The reason that they do this kind of terrorism is because of Islam. This is how the people are feeling about the terrorist attacks in their countries. In this figure, we can see that the countries that suffer most are America and Britain. This is because they have many people, and many people are in danger because of the weapons and bomb attacks.

Figures 6, 7, and 8 show that in X-axis countries are given, and in Y-axis, it shows that the number of people has killed in the terrorist attacks is showing. The figure clearly shows that the death rate also varies according to the terrorist attack in different countries.

Terrorist acts in the Middle East and northern Africa have been seen to have fatal consequences. The Middle East and North Africa are seen to be places of severe terrorist attacks. If you look at the graphics, Iraq, Afghanistan, and Pakistan are the most broken countries. All of these countries are Muslim countries. In addition, even

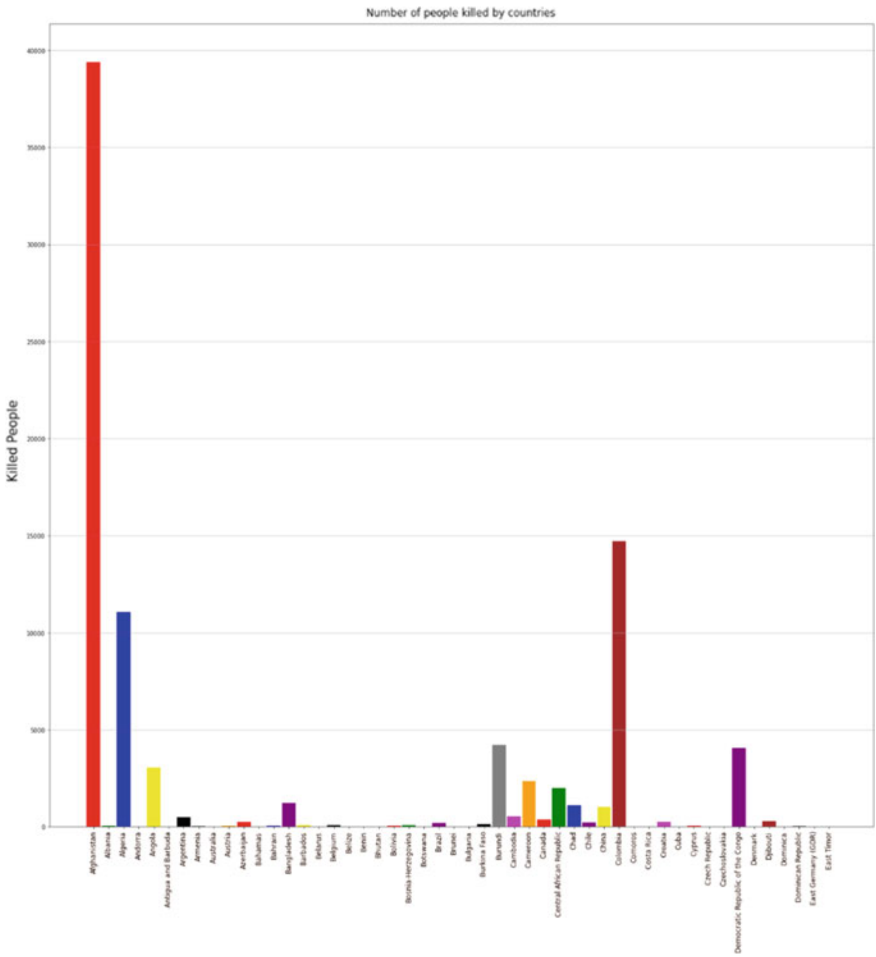


Fig. 6 Fatality rate of different countries in terrorist attacks over a decades

though there is a perception that Muslims are supporters of Terrorism, Muslims are the most damaged by terrorist attacks.

4.4 Discussion

Furthermore, the United States (USA) is the most damaged country. This happens because it is an important country, and it has been perceived as the leader of the wars against terrorism. Furthermore, the USA was thought to be where the global fight against terrorism was taking place. In addition, the Arab countries and other Muslim countries were considered to be allies in the worldwide fight against terrorism.

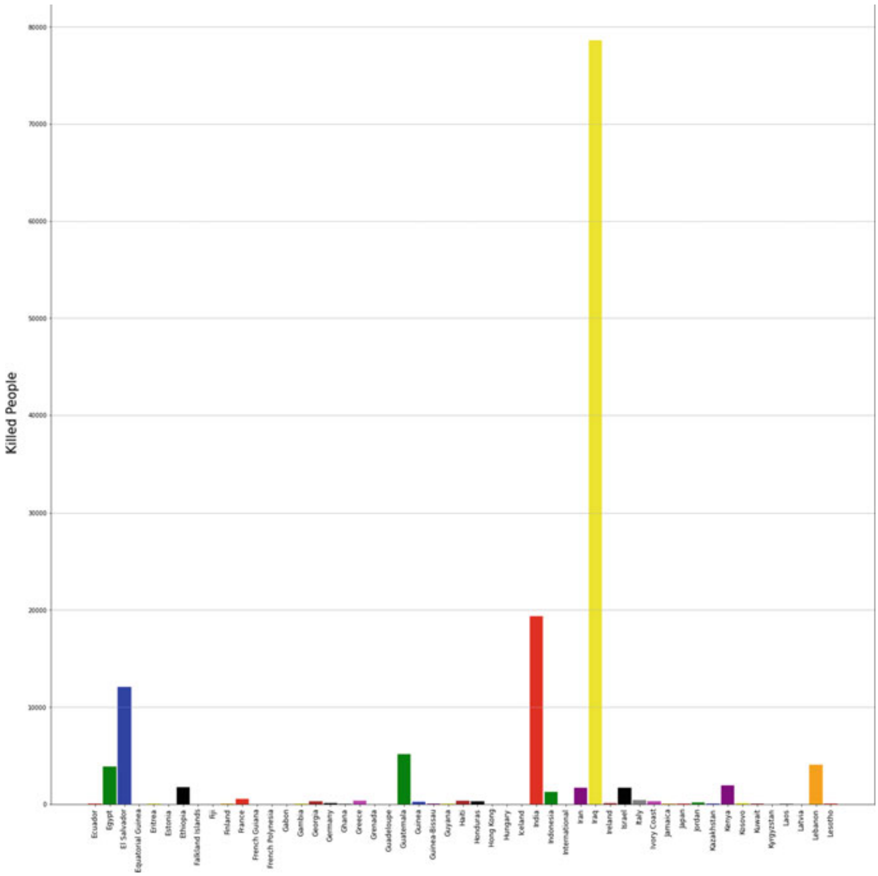


Fig. 7 Mortality rate of different countries in terrorist attacks (within home region terrorist cases)

However, ISIS managed to turn these allies into the enemy. After the destruction caused by ISIS, the USA has left the most damaged country in the world. Therefore, there is a global war against terrorism. It is a war where all the nations have lost their citizens, resources, and land fight. There are several other reasons why the Middle East and North Africa are the most damaged places by the war against terrorism. For instance, the Middle East and North Africa were hit hard by ISIS because they highly depended on Middle Eastern oil.

Additionally, these countries were under pressure to join the fight against ISIS. Furthermore, due to security reasons, ISIS could take power in some regions of these countries. There are different types of terrorist attacks that can be seen. In addition, there are many terrorist acts in some of these countries. Terrorist acts can occur because of several factors such as terrorism training, a lack of proper education, or religious differences. In addition, terrorist acts can be seen in Syria, Iraq, Palestine, Yemen, Libya, Egypt, Mali, Algeria, and Morocco. In addition, terrorist attacks can

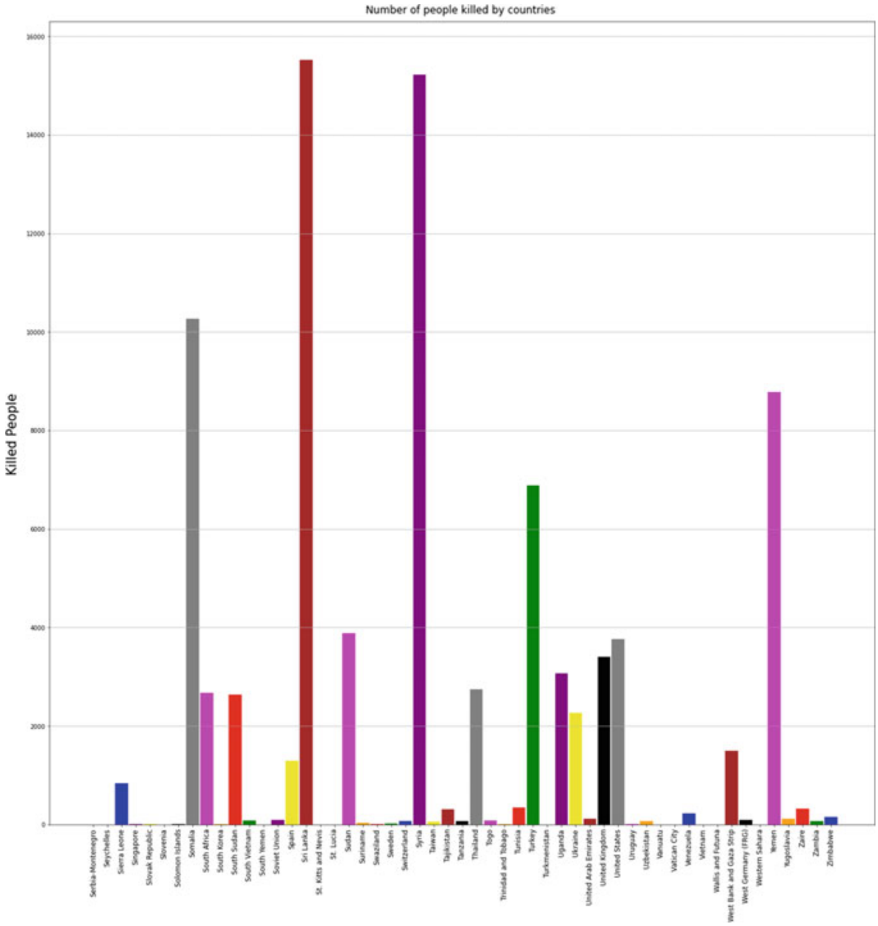


Fig. 8 Fatality rate of terrorists in different countries using some successful operation

be divided into two different categories. They are the acts that happened when some terrorist was trained and those that occurred because of a lack of education.

5 Conclusion

This article gives you an overview of the causes of terrorism and the significant zones or the hot zones of terrorism, why it is considered a danger for the people, and what our motive to deal with it should be. Here, we tried to propose some solutions for the issues. While considering all the aspects, we have attempted to analyze the Global terrorism dataset using various features like we have tried to figure out the graph using different data analysis approach. Summing up all the things first part of this article deals with the essential aspects of terrorism, significant causes, and all the factors contributing to it, the second part deals with the different data analysis approach using the dataset and the mathematical formulation behind it, and we have concluded the article in the third part.

Future Work The above article presents an overview of analyzing the terrorist dataset with the help of specific machine learning techniques. In the coming times, we plan to explore and train this dataset using some neural network model and identify the hot zones about the attacks and what is behind the curtains. We would like to see the dataset used with the help of different machine learning models to observe how well they work and how accurate they can predict the occurrence of attacks, location, targets, and so on.

References

1. Kis-Katos K, Liebert H, Schulze GG (2011) On the origin of domestic and international terrorism. Eur J Polit Econ 27:S17–S36
2. Mullins WC (1997) A sourcebook on domestic and international terrorism: an analysis of issues, organizations, tactics, and responses. CC Thomas
3. Singer G, Golan M, Identification of subgroups of terror attacks with shared characteristics for the purpose of preventing mass-casualty attacks: a data-mining approach
4. Ruby CL (2002) The definition of terrorism. In: Analyses of Social Issues and Public Policy (ASAP)
5. Pravallika CR, A machine learning approach for enhancing defence against global terrorism
6. Dansana D, Kumar R, Das Adhikari J et al (2020) Global forecasting confirmed and fatal cases of COVID-19 outbreak using autoregressive integrated moving average model. Front Public Health 8:580327. Published 29 Oct 2020.https://doi.org/10.3389/fpubh.2020.580327
7. Dansana D, Kumar R, Parida A, Sharma R, Das Adhikari J, Van Le H, Thai Pham B, Kant Singh K, Pradhan B (2021) Using susceptible-exposed-infectious-recovered model to forecast coronavirus outbreak. Comput Mater Continua 67(2):1595–1612. https://doi.org/10.32604/cmc.2021.012646
8. Dansana D, Adhikari JD, Mohapatra M, Sahoo S (2020) An approach to analyse and forecast social media data using machine learning and data analysis. In: 2020 International conference on computer science, engineering and applications (ICCSEA). https://doi.org/10.1109/iccsea49143.2020.9132895
9. White JR, Clear T (2002) Terrorism: an introduction. Wadsworth Thomson Learning, Stanford, CA
10. Crenshaw M (ed) Terrorism in context. Penn State Press

Intelligent Multiple Diseases Prediction System Using Machine Learning Algorithm

Sudheer Babu, Dodala Anil Kumar, and Kotha Siva Krishna

Abstract As a result of their surroundings and lifestyle choices, people nowadays suffer from a wide range of ailments. As a result, predicting illness at an early stage is crucial. Doctors, on the other hand, struggle to make accurate diagnoses based solely on symptoms. The most challenging task is predicting sickness properly. Machine learning plays a key part in forecasting in order to complete this difficult task. To tackle this challenge, machine learning plays a key role in illness prediction. Medical research creates a vast amount of data every year. Early patient care has benefitted from effective medical data analysis because of the rising quantity of data growth in the medical and healthcare professions. In data mining, disease data is utilised to identify hidden patterns in huge volumes of medical data. Based on the patient's symptoms, we created a broad disease prediction. Machine learning algorithms like ANFIS and CNN are used to properly predict sickness (adaptive network-based fuzzy inference system). The collection of illness symptoms is necessary for disease prediction. For an accurate prognosis, this general illness prediction takes into account the person's lifestyle and medical history. When it comes to illness prediction, ANFIS outperforms CNN by a wide margin (96.7%). ANFIS, on the other hand, does not require as much time or memory to train and test because it does not use the UCI repository dataset. There are several libraries and header files included with the Anaconda (Jupyter) notebook that make Python programming more precise and accurate.

Keywords CNN · ANFIS · Machine learning · Disease prediction

S. Babu (✉) · D. Anil Kumar · K. Siva Krishna
GIET University, Gunupur, Odisha 765022, India
e-mail: sudheerpunuri@giet.edu

D. Anil Kumar
e-mail: danil@giet.edu

K. Siva Krishna
e-mail: ksivakrishna@giet.edu

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_55

1 Introduction

Classification algorithms, as opposed to individual classifiers, are extensively used in the medical business to categorise data into various classes based on given criteria. Diabetes occurs when the body's capacity to create insulin has been hindered, leading to improper carbohydrate metabolism and elevated blood glucose levels. High blood sugar levels are a symptom of diabetes, which is characterised by this. Elevated blood sugar levels are accompanied by symptoms such as increased thirst, hunger and frequent urination. Diabetic complications might arise if the disease is not properly managed. Two of the most significant consequences are diabetic ketoacidosis and nonketotic hyperosmolar coma [1]. Diabetes is a major health problem in which the quantity of sugar in the body cannot be managed. Diabetes is influenced by a number of factors, including height, weight, genetic factors, and insulin, but the sugar concentration is the most important aspect to consider. Complications can only be avoided if they are identified as soon as feasible [2].

An excessively high blood glucose level, often known as blood sugar, is indicative of diabetes mellitus. We get our energy from glucose, a key component of carbohydrates, which is found in our bloodstreams. Insulin, a hormone produced by the pancreas, is responsible for regulating blood sugar levels. Glucose from meals is absorbed and converted into energy by beta cells. There are times when the human body does not create enough insulin, or there is no insulin at all in the human body. It is still possible to control type 2 diabetes by early identification and diagnosis as well as adequate medical treatment and easy lifestyle adjustments [3].

To save lives, doctors are increasingly relying on data mining tools to examine medical databases to help them detect disease at an early stage. For generations, doctors have depended on medical data to diagnose diabetes, which affects a huge portion of the world's population. Patients' age, gender, symptoms, insulin levels, blood pressure, blood glucose levels and weight are included in these databases, which are updated on a regular basis [4]. Data mining techniques are used to extract and use just the most relevant data to aid in correct diagnosis because of the vast quantity of data available. It is possible to employ machine learning methods like decision trees, Naïve Bayes and support vector machines to aid in the right identification of disease-related data [5].

Specifically, this work focuses on the ANFIS machine learning classification algorithm's UCI Diabetes Training Dataset, which was derived from the dataset. This strategy was employed instead of analysing all of the diabetes-related features in the dataset to find those that had the best chance of predicting the disease's incidence. For diabetic drug planning purposes, ANFIS is a more effective and trustworthy approach than other methods, and the outcomes achieved for each algorithm are evaluated by comparing accuracy measures.

For the rest of this essay, it is arranged as follows. Section 2 analyses other literature surveys, Sect. 3 explains the recommended strategy, and Sect. 4 summarises the findings. In this section, we summarise the study's findings.

2 Literature Survey

Data analysis and prediction of future behaviour using datasets are beneficial in a variety of resources, including human health. Human resources are one of the most important and everlasting components of this planet. As digital technology expands in every sector, it is also expanding in medical facilities to detect and predict diseases such as diabetes. A specific number of studies have been conducted by analysing and predicting future diseases based on datasets provided by patient's health records. We discovered some of the best methods for prediction and analysis through additional records and work on medical technology.

In Peek et al. [6], the WEKA data mining classification tool was judged to be one of the best. J48 has a 74.28% accuracy rate on the data set for breast cancer, whereas SMO has a 76.80% accuracy rate on the database of data about diabetes, according to them. They used a selection approach on a variety of algorithms and methodologies to improve the accuracy of different algorithms [7]. After feature selection, they get 82.22 per cent for Decision Tree, 82.56 per cent for Logistic Regression, 84.17 per cent for Random Forest, 84.24 per cent for Naïve Bayes and 84.85 per cent for Logistic Regression SVM. In [8], they used k-nearest neighbour and a genetic algorithm to examine large data sets for the prediction of coronary heart disease. They came to the conclusion that the provided method allows them to anticipate diseases at an early stage and that it helps to the progress of medicine by making it easier for physicians to assess and predict diseases. In [9], they developed a cardiovascular disease prediction algorithm based on data from wearable devices. Using the Logistic Regression Model, they were able to attain an accuracy of 87 per cent. Furthermore, measuring under the ROC curve results in an AUC value of 80%. They used ML classifiers to do a number of comparative evaluations for the prediction of heart and hepatitis disorders in [10]. On the same database, they utilised six different classifiers: LR, DT, NB, KNN, SVM and RF. After running the whole algorithm on the same dataset, they discovered that the recommended technique outperformed all other classifiers in terms of performance and prediction (Table 1).

Because chronic illnesses progress slowly, getting a thorough diagnosis and treatment as soon as feasible is critical. As a result, it is critical to create a decision model that can aid in the identification of chronic diseases and the prediction of patient outcomes in the near term. In spite of the fact that there are several ways, the current study concentrates on one. On machine learning prediction models used in CD diagnosis, emphasising the significance of this research, we conducted a complete review of the literature on numerous state-of-the-art predictive models in this study, and our main contribution is a model classification suggestion based on comparative model analysis. The results of this review paper will be able to identify possible outcomes that will improve the quality of patient data and the analysis of specific things associated with machine learning algorithms in medical care when compared to standard data analysis approaches, as compared to standard data analysis approaches. The results of this review paper will be able to identify possible outcomes that will improve the quality of patient data and the analysis of specific things associated

Table 1 Related work on various approaches of analysis and predicting future diseases

Sl. no.	References	Techniques	Drawbacks
1	[11]	They show clinical factors that, when paired with the right data cleaning approach, increase the accuracy of disease progression and survival rate estimates in Amyotrophic Lateral Sclerosis	However, experimental results show that the proposed methods did not perform well in terms of slope and survival predictions
2	[12]	In terms of accuracy, the Support Vector Machine was shown to be the most accurate in sickness prediction	In general, the Support Vector Machine (SVM) outperforms the Naïve Bayes (NB), the Random Forest (RF) and the Simple Cart (SC). As a result, the efficacy of the suggested model is clearly demonstrated to be insufficient for future modes
3	[13]	In this paper, two data mining methods, Naïve Bayes and the J48 algorithm, are compared for accuracy and performance on training medical datasets	The conclusion derived from the J48 and Naïve Bayes evaluation results was that Naïve Bayes results did not satisfy the best accuracy of prediction of analysis on the datasets
4	[14]	It consists of a disease prediction system that employs machine learning and artificial neural networks concepts	It does not provide complete data functionalities from datasets
5	[15]	A vast quantity of patient data is analysed using data mining and machine learning algorithms, which results in the application of big data technologies and tools	It concludes that we are unable to predict all diseases at an early stage using this predicted system and that this system cannot be used in the medical field for physicians to easily analyse and predict disease
6	[16]	This study presented a generic illness prediction based on patient symptoms using standard methods	It is difficult to analyse and extract diabetes disease data
7	[17]	This study offered broad illness prediction based on patient symptoms using KNN and CNN	Although KNN and CNN algorithms are used to categorise patient data, and they do not produce good predictions for clinicians to be satisfied
8	[18]	This study focuses on the selection of techniques and algorithms for experimentation analysis using multiple heart disease datasets	A decision tree has an accuracy of 82.22 per cent, a logistic regression has an accuracy of 82.56 per cent, a random forest has an accuracy of 84.17 per cent, a Naïve Bayes has an accuracy of 84.24 per cent, and a logistic regression SVM has an accuracy of 84.85 per cent

with machine learning algorithms in medical care when compared to standard data analysis approaches.

3 Proposed Methodology

Identifying classification algorithms that will most reliably predict the existence of diabetes in patients is the major objective of this article, which will allow for early identification and treatment. Figure 1 depicts the model diagram for the proposed work. It is possible to predict more than one disease at a time using a multi disease model prediction. As a result, users do not need to navigate through a plethora of models in order to predict diseases. It will save time, and by predicting multiple diseases at once, the mortality rate may be reduced.

3.1 Preprocessing

Certain preprocessing techniques on the network inputs and targets may improve the efficiency of neural network training. The network input processing functions transform inputs into a network-friendly format. The data preparation procedure for raw inputs is heavily influenced by the normalisation process. Without this normalisation, training the neural networks would have been incredibly slow. There are several different types of data normalisation. It may be used to scale data for each

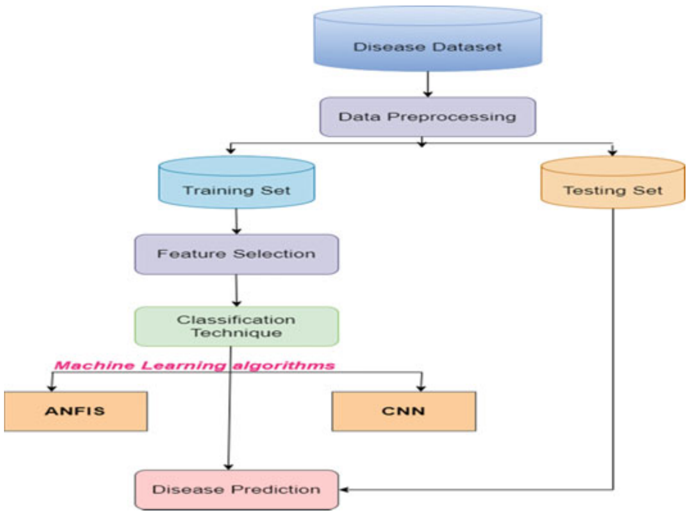


Fig. 1 Proposed system flowchart

input feature in the same range of values, removing bias from one feature to the next inside the neural network. Data normalisation can shorten training times by starting each feature's training on the same scale. A wide variety of sizes [19] can be used for modelling applications that need a wide range of inputs.

Feature Selection.

A prominent data mining approach is Feature Selection, also known as Variable Selection [20], for removing unnecessary and superfluous characteristics from any dataset [21]. This method also increases data understanding, provides for improved data visualisation, decreases learning algorithm training time and improves prediction performance. A rapid and efficient feature identification algorithm was utilised. The use of symmetric uncertainty (SU) led to the identification of an optimal threshold value. The least spanning tree was built when symmetric uncertainty (SU) was applied. On the basis of the accuracy of categorization and the proportion of characteristics included chosen, the results of the recommended algorithm were compared to those of other algorithms such as FAST, FCBF, Relief and CFS, and it was determined that Modified FAST was the best among them [22].

Classification.

CNN.

The Multilayer Perceptron (MLP) is a form of convolutional neural network (CNN) that is a subclass of the convolutional neural network (CNN) (MLP). Their similarities to neural networks include the following characteristics: weakening and biases must be learned by neurons in order to work. There are several ways in which each neuron receives information. Afterwards, if necessary, a dot product operation and a nonlinearity function are used. As a result, CNNs were initially utilised in image processing for the purpose of transforming raw picture pixels into class scores. Among CNN's three major levels, there are three distinct sections. Layers include a convolutional, a pooling; additionally, there is a completely linked layer with a corrected linear activation function. We make use of convolution 1D layers, pooling 1D layers and other techniques and fully connected layers because our application is focussed on assessing one-dimensional input. [23] In this case, CNN organises the one-dimensional time series data sequentially.

In order to create an entirely new feature map, the following procedure is followed:

$$hlifm = \tanh(Wfmxi : i + f - 1 + b) \tag{1}$$

The filter *hl* is employed to each set of features f in the input data defined by $\{x1: f,\ x2: f + 1, \ldots, xn - f + 1\}$ so as to generate a feature map as $hl = [hl1, hl2, \ldots, hln - f + 1]$.

There is a so-called tmax function included in this fully linked layer, which estimates the probability distribution for each class. As a result, the classes that make up the CNN network's final output will be computed by the fully connected layer (FC). The CNN's architecture is therefore INPUT-CONV-POOL-FC (Fig. 2).

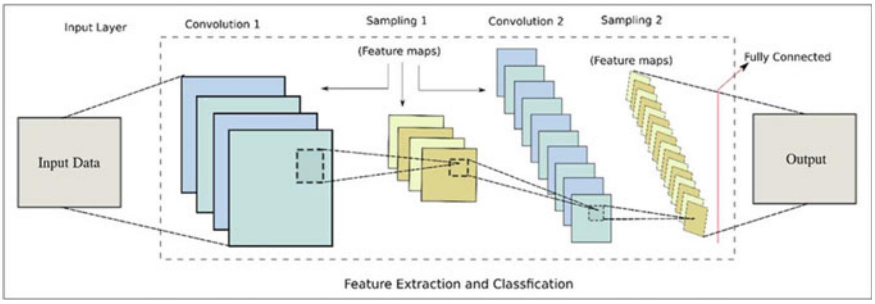


Fig. 2 CNN classification architecture

ANFIS.

Systems that use neuro-fuzzy inference and artificial neural networks can be divided into two categories. There is fuzzy inference in the first situation, but not in the second. Mamdani and Takagi–Sugeno fuzzy reasoning are two of the second types of techniques that have been extensively studied as a way to integrate artificial neural networks. Finally, Mamdani's fuzzy set and Takagi–first-order Sugeno's linear equations of input variables are at their final states after this rule. The Takagi–Sugeno system, for example, is computationally efficient, adaptable to new rules, and has the benefit of keeping output space continuity when paired with an artificial neural network optimization strategy. ANFIS used the Takagi–Sugeno approach to build and infer rules because of the reasons outlined above (Fig. 3).

Topalovic et al. [24] presented the ANFIS model as a kind of neuro-fuzzy inference system. The least square approach (LSE) and back propagation (BP) methods, used by the ANFIS model, allow it to obtain the most accurate estimates of the parameters contained in the membership functions and output. In order to show the

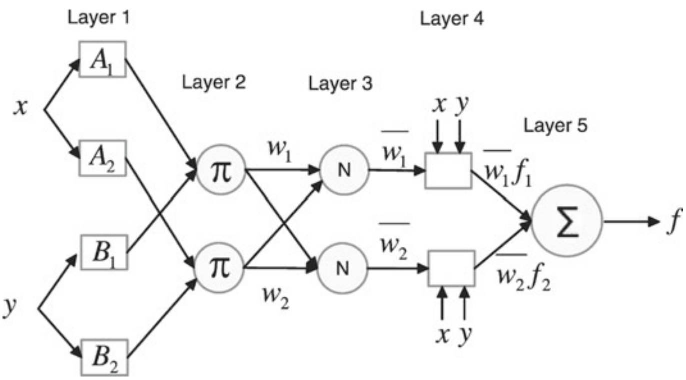


Fig. 3 ANFIS classification architecture

fuzzy inference process, we use a model consisting of n Takagi–Sugeno rules, each with two inputs and the first-order linear equation result.

4 Results and Discussion

The suggested technique is validated using data from the UCI Repository's Diabetes Dataset (UCI) [25]. The medical records of 768 female patients are included in this collection. A set of symptoms reported by the user is put into the illness prediction algorithms. On the platform, the symptoms may be chosen from a drop-down menu, and the algorithms will analyse them on the backend. The user-selected symptoms are then analysed by each algorithm, with the mean accuracy being the result. For such methods, the result is then utilised as weights in the ensemble classifier. The assessment measures employed are Accuracy, Specificity, Precision and Recall [26].

Accuracy is a metric that compares the number of true-positive (TP) and true-negative (TN) test pictures to the total number of test photos.

$$\text{Accuracy} = \text{TP} + \text{TNTP} + \text{TN} + \text{FP} + \text{FN} \tag{2}$$

(i) *Precision refers to the estimate of genuine positives to the total value of true positives and false positives.* It is written as Eq. (3)

$$\text{Precision} = \text{TP}(\text{TP} + \text{FP}) \tag{3}$$

(ii) It is the estimation analysis of the true-positive rate to the total value of the true-positive and false-negative rates, as previously stated. It is written as Eq. (4).

$$\text{Recall} = \text{TP}(\text{TP} + \text{FN}) \tag{4}$$

(iii) Specificity: It compares the true-negative analysis to the total value of false positives and true negatives. It is written as Eq. (5)

$$\text{Specificity} = \text{TN}(\text{FP} + \text{TN}) \tag{5}$$

True positives are denoted by the letters TP, true negatives by the letters TN, while actual positives and negatives are denoted by the letters P and T, respectively. A effective predictor will have a high sensitivity, low specificity and a high accuracy.

Table 2 shows the properties of diabetes dataset's attributes. Preg, plas, pres, skin and age are the five characteristics in the Diabetes dataset (Tables 3, 4, 5, 6 and 7).

Table 2 displays the diabetes attributes dataset's description

Sl. No.	Attribute name	Description
1	Preg	The number of times a pregnant woman's plasma glucose concentration was measured was calculated as
2	Plas	During a two-hour period, participants took part in an oral glucose tolerance test
3	Pres	Blood pressure in the diastole (mmHg)
4	Age	Age(in years)

Table 3 Summarization of prediction techniques with performance for preg

Prediction techniques	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)
SVM [27]	0.69	0.67	0.62	0.60
KNN [28]	0.71	0.70	0.68	0.61
Decision Tree [29]	0.80	0.78	0.72	0.69
CNN [30]	0.89	0.87	0.83	0.70
Proposed ANFIS	0.97	0.95	0.94	0.92

Table 4 Summarization of prediction techniques with performance for plas

Prediction techniques	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)
SVM [27]	0.74	0.71	0.68	0.63
KNN [28]	0.75	0.72	0.66	0.60
Decision Tree [29]	0.76	0.74	0.69	0.61
CNN [30]	0.80	0.78	0.71	0.67
Proposed ANFIS	0.96	0.94	0.93	0.91

Table 5 Summarization of prediction techniques with performance for pres

Prediction techniques	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)
SVM [27]	0.76	0.73	0.70	0.67
KNN [28]	0.77	0.74	0.70	0.67
Decision Tree [29]	0.80	0.77	0.73	0.69
CNN [30]	0.84	0.80	0.78	0.75
Proposed ANFIS	0.95	0.94	0.93	0.92

Table 6 Summarization of prediction techniques with performance for age

Prediction techniques	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)
SVM [27]	0.66	0.63	0.60	0.58
KNN [28]	0.70	0.69	0.64	0.60
Decision Tree [29]	0.75	0.70	0.68	0.65
CNN [30]	0.80	0.77	0.74	0.70
Proposed ANFIS	0.95	0.93	0.92	0.91

Table 7 Summarization of CPU time test for all classification algorithms

Prediction techniques	Running time (ms)
SVM	5.63
KNN	3.75
Decision Tree	2.19
CNN	0.94
Proposed ANFIS	0.36

5 Conclusion

The great majority of the world's population is affected by diabetes. Diabetes is a disease that can only be controlled, not cured. Diabetes that is not well controlled can cause complications that can lead to a range of other chronic conditions. As a result, it is vital to have early identification, successful treatment and adequate diabetes management. Diabetic nerve disorders can impact the heart and, as a result, the heart rate. In this article, we employ UCI data (extracted from a variety of electric signals) to accurately identify diabetes. This is the first study that we are aware of that uses deep learning to identify diabetes using UCI data. The ANFIS classification system has a 96.7 per cent accuracy, which is the highest accuracy in automated diabetes identification using UCI to date. For benchmark comparison with the proposed ANFIS classification, explicit feature extraction and the usage of classical classifiers are necessary. With the suggested approach, clinicians will be able to diagnose diabetes more precisely. As mentioned in the Results section, a bigger input dataset than the one utilised in this study can be fed into the suggested architecture to see whether accuracy can be improved further.

References

1. Hamet P, Tremblay J (2017) Artificial intelligence in medicine. Metabolism 69:S36–S40
2. Johnson KW, Soto JT, Glicksberg BS, Shameer K, Miotto R, Ali M, Dudley JT (2018) Artificial intelligence in cardiology. J Am Coll Cardiol 71:2668–2679
3. Buch VH, Ahmed I, Maruthappu M (2018) Artificial intelligence in medicine: current trends and future possibilities. Br J Gen Pract 68:143–144
4. Kotsiantis SB, Zaharakis I, Pintelas P (2007) Supervised machine learning: a review of classification techniques. Emerg Artif Intell Appl Comput Eng 160:3–24
5. Deo RC (2015) Machine learning in medicine. Circulation 132:1920–1930
6. Peek N, Combi C, Marin R, Bellazzi R (2015) Thirty years of artificial intelligence in medicine (AIME) conferences: a review of research themes. Artif Intell Med 65:61–73
7. Battineni G, Sagaro GG, Nalini C, Amenta F, Tayebati SK (2019) Comparative machine-learning approach: a follow-up study on type 2 diabetes predictions by cross-validation methods. Machines 7:74
8. Lo Y-C, Rensi S, Torng W, Altman RB (2018) Machine learning in chemoinformatics and drug discovery. Drug Discov Today 23:1538–1546

9. Napolitano G, Marshall A, Hamilton P, Gavin AT (2016) Machine learning classification of surgical pathology reports and chunk recognition for information extraction noise reduction. Artif Intell Med 70:77–83
10. Polat H, Mehr HD, Cetin A (2017) Diagnosis of chronic kidney disease based on support vector machine by feature selection methods. J Med Syst 41:55
11. Eslamizadeh G, Barati R (2017) Heart murmur detection based on wavelet transformation and a synergy between artificial neural network and modified neighbor annealing methods. Artif Intell Med 78:23–40
12. Martinez D, Pitson G, Mackinlay A, Cavedon L (2014) Cross-hospital portability of information extraction of cancer staging information. Artif Intell Med 62:11–21
13. Wells G, Shea B, O'Connell D, Peterson J (2000) The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute: Ottawa, ON, Canada
14. PRISMA (2015) PRISMA—Transparent reporting of systematic reviews and meta-analyses. Ottawa Hospital Research Institute, Ottawa, ON, Canada
15. Chen Y, Luo Y, Huang W, Hu D, Zheng R-Q, Cong S-Z, Meng F, Yang H, Lin H, Sun Y et al (2017) Machine-learning-based classification of real-time tissue elastography for hepatic fibrosis in patients with chronic hepatitis B. Comput Boil Med 89:18–23
16. Shousha HI, Awad AH, Omran D, Elnegouly MM, Mabrouk M (2018) Data mining and machine learning algorithms using IL28B genotype and biochemical markers best predicted advanced liver fibrosis in chronic hepatitis C. Jpn J Infect Dis 71:51–57
17. Zhou L-Q, Wang J-Y, Yu S-Y, Wu G-G, Wei Q, Deng Y-B, Wu X-L, Cui X-W, Dietrich CF (2019) Artificial intelligence in medical imaging of the liver. World J Gastroenterol 25:672–682
18. Zhou W, Ma Y, Zhang J, Hu J, Zhang M, Wang Y, Li Y, Wu L, Pan Y, Zhang Y et al (2017) Predictive model for inflammation grades of chronic hepatitis B: large-scale analysis of clinical parameters and gene expressions. Liver Int 37:1632–1641
19. Mcheick H, Saleh L, Ajami H, Mili H (2017) Context relevant prediction model for COPD domain using bayesian belief network. Sensors 17:1486
20. Shah SA, Velardo C, Farmer AJ, Tarassenko L, Kim B, Barber V (2017) Exacerbations in chronic obstructive pulmonary disease: identification and prediction using a digital health system. J Med Int Res 19:e69
21. Granero MAF, Sanchez-Morillo D, Leon-Jimenez A (2015) Computerised analysis of telemonitored respiratory sounds for predicting acute exacerbations of COPD. Sensors 15:26978–26996
22. Olivera AR, Roesler V, Iochpe C, Schmidt MI, Vigo Á, Barreto SM, Duncan BB (2017) Comparação de algoritmos de aprendizagem de máquina para construir um modelo preditivo para detecção de diabetes não diagnosticada—ELSA-Brasil: Estudo de acurácia. Sao Paulo Med J 135:234–246
23. Liu X, Li N, Lv L, Fu Y, Cheng C, Wang C, Ye Y, Li S, Lou T (2017) Improving precision of glomerular filtration rate estimating model by ensemble learning. J Transl Med 15:231
24. Topalovic M, Laval S, Aerts JM, Troosters T, Decramer M, Janssens W (2017) Belgian pulmonary function study investigators. Automated interpretation of pulmonary function tests in adults with respiratory complaints. Respiration 93:170–178
25. Battineni G, Chintalapudi N, Amenta F (2019) Machine learning in medicine: performance calculation of dementia prediction by support vector machines (SVM). Inform Med Unlocked 16:100200
26. Griffis JC, Allendorfer JB, Szaflarski JP (2015) Voxel-based Gaussian naïve Bayes classification of ischemic stroke lesions in individual T1-weighted MRI scans. J Neurosci Methods 257:97–108
27. Dinga R, Marquand AF, Veltman DJ, Beekman ATF, Schoevers RA, Van Hemert AM, Penninx BWJH, Schmaal L (2018) Predicting the naturalistic course of depression from a wide range of clinical, psychological, and biological data: a machine learning approach. Transl Psychiatry 8:241
28. Pekkala T, Hall A, Lötjönen J, Mattila J, Soininen H, Ngandu T, Laatikainen T, Kivipelto M, Solomon A (2016) Development of a late-life dementia prediction index with supervised machine learning in the population-based CAIDE study. J Alzheimers Dis 55:1055–1067

29. Kuo SH, Lin CY, Wang J, Sims PA, Pan MK, Liou JY, Faust PL (2017) Climbing fiber-Purkinje cell synaptic pathology in tremor and cerebellar degenerative diseases. Acta Neuropathol 133:121–138
30. Oakden-Rayner L, Carneiro G, Bessen T, Nascimento JC, Bradley A, Palmer LJ (2017) Precision radiology: predicting longevity using feature engineering and deep learning methods in a radiomics framework. Sci Rep 7:1648

Detecting Brain Tumors in Medical Image Technology Using Machine Learning

Bhaskar Mekala and P. Kiran Kumar Reddy

Abstract With an expansion in the demand for automated medical imaging, the field is getting importance, fast, reliable and efficient diagnosis which can provide insight to the picture image better than human eyes. Brain tumor is the second leading cause of cancer-related deaths in men age 20–40 and 5th leading cause cancer among women in the same group. A diagnosis of tumor is a very important part in its treatment. Identification of a tumor is very important part in its treatment. To obtain the background, this paper covers noise elimination and image sharpening and also morphological functions, erosion and dilation. Plotting contour and c-label of the tumor and its boundary provides us identifying the size, shape and position of the tumor, it helps the medical employee as well as the patient to understand the seriousness of the tumor with the help of different labeling for different levels of elevation.

Keywords Brain tumor · Machine learning · CT image · MRI scanners · Plotting contour · Tumor detection

1 Introduction

The main impetus of this undertaking is to create a directforword climate where clinical staff and patient can work in complete participation to accomplish better outcomes. This directforword climate will assist the patient with having a sense of safety as they will comprehend the therapy interaction decision, which will assist the clinical staff with taking care of the circumstances in a request giving them more opportunity to think and work.

B. Mekala
GIET University, Gunupur, Odisha, India
e-mail: Bhaskar.mekala@mriet.ac.in

P. Kiran Kumar Reddy (✉)
MRIET-JNTU, Hyderabad, India
e-mail: kiran.penubaka@gmail.com

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_56

A brain tumor [1, 2] is characterized by an abnormal growth of cells inside the mind or a specific spinal column. A few growths can be harmful in this manner they should be identified and relieved on schedule. The specific reason for cerebrum growths is not clear nor is accurate arrangement of indications characterized; consequently, individuals might be experiencing it without understanding the risk. Essential mind growth can be either dangerous or harmless disease cells [3].

Cerebrum growth happened when the cells were separating and developing strangely. It is having all the earmarks of being a strong mass when it determined to have indicative clinical imaging methods. There are two kinds of cerebrum cancer which is essential mind growth and metastatic brain cancer. Essential brain growth is the condition when the cancer is shaped in the mind and would in general, remains there while the metastatic brain cancer is the growth that is framed somewhere else in the body and spread through the mind [4].

The indications of a tumor are influenced by the tumor's location, size, and type of the tumor and it occurs when a tumor pushes the surrounding cells and exerts pressure upon them. Besides, it is also occurring when the tumor blocks the fluid that flows throughout the brain. The normal side effects are having cerebral pain, queasiness and regurgitating and having issue in adjusting and strolling. Cerebrum growth can be distinguished by the indicative imaging modalities, for example, CT output and MRI. Both of the modalities enjoy benefits in distinguishing relying upon the area type and the reason for assessment required. In this paper, we like to utilize the CT pictures since it is difficult to look and gives out exact calcification and unfamiliar mass area [4].

The CT image acquired from the CT machine gives two-dimensional cross section of brain. However, the image acquired did not extract the tumor from the image. Thus, the image processing is needed to determine the severity of the tumor depends on the size [4, 5].

Selecting CT images upon MRI images is as follows:

1. CT is much faster than MRI, making it the study of choice in cases of trauma and other acute neurological emergencies. CT can be obtained at considerably less cost than MRI.
2. CT image cost is less than MRI image.
3. CT is less sensitive to patient motion during the examination.
4. CT can be performed at no risk to the patient with implantable medical devices, such as cardiac pacemakers, ferromagnetic vascular clips and nerve stimulators.
5. The imaging can be performed much more rapidly, so CT may be easier to perform in claustrophobic or very heavy patients.

The focus of this project is CT brain images tumor extraction and its representation in simpler form such that it is understandable by everyone. Humans tend to understand colored images better than black and white images; thus, we are using colors to make the representation simpler enough to be understood by the patient along with the medical staff. Contour plot and c-label of tumor and its boundary is programmed to give 3D visualization from 2D image using different colors for different levels of intensity. A user-friendly GUI is also created which helps medical staff to attain the above objective without getting into the code.

2 Literature Survey

This part surveys the current techniques dependent on brain tumor division. This part additionally clarifies the benefits and restrictions of existing techniques.

As per one more author, Zeineldin team using the methodology using a deep neural network framework [6] for automatic brain tumor segmentation was developed using magnetic resonance, fluid attenuation inversion to restore images and this paper has some advantages; the CNN encoder part is responsible for spatial information extraction then insert the resulting semantic map into the decoder section to obtain a full-resolution probability map. But this paper has few limitations like, due to the percentage of overlap between ground truth segmentation with prediction, the developed method showed lower performance. In contrast, no false positives indicate high specificity and sensitivity, which may not accurately reflect actual performance.

As per Zhou, Z., He, Z., Jia [1] followed the methodology set up 3D completely convolution neural organization with atrous convolution trademark pyramid for cerebrum cancer division through MRI pictures. Atrous convolution is an option for the down examining layer. It expands the responsive field while keeps up with the spatial element of element maps.

This paper has mostly focused a 3D completely associated conditional random field is made as a post-preparing venture for the organization's yield to get underlying division of both the looks and spatial consistency. Zhou and group examined restrictions, and the fundamental impediment is the open field of four-progression pyramid simply too enormous to even think about joining up with the foundation and in this way the division of the non-upgrading center contrasted and different injuries include a lower-DSC esteem. Since the non-improving center is commented on the grounds that the excess a piece of the gross cancer center.

As per some more authors, Chen, S., Ding, C. and Liu presented a paper title "Dual-force convolution neural networks for accurate brain tumor segmentation" [7] actually it is the methodology for accurate brain tumor segmentation.

This paper has following advantages:

First advantage of this paper is the multi-layer perception (MLP)-based method for post-processing, which may refine the prediction results of CNNs and further promote the segmentation performance. And this advantage is compared with existing method, and it is light and straightforward to use in practice.

Author Chen and team discussed about the limitations of this concept, and the improvement is not significant and there is even subtle decline in dice score for the enhancing tumor. We argue that this is often because the standard of the learnt multi-level features is not optimal.

As per authors Narmatha, C., Eljack and group utilizing the procedure using paper title "Carried out a fuzzy brain-storm optimization (FBSO) calculation for division and characterization of mind growths," [4] this paper basically focusing on a 3D completely associated conditional random field is developed as a post-preparing venture for the organization's yield to acquire primary division of both the appearance and spatial consistency. It contains not many limitations; the FBSO strategy did not think about the significant components, which devours just 93.85% of exactness. Subsequently, the created strategy requires successful component extraction procedures.

According to Ismail, S. A. A., Mohammed, A. and Hefny, using the methodology, developed an enhanced technique using the residual network (ResNet) to classify the tumors in brain [8]. And this paper mainly concentrates on the ResNet which allows the training of extremely deep neural networks by shortcut connection concept; therefore, solving the problem of degrading accuracy and vanishing gradient which happened when models got deeper. This paper has few limitations: The method achieved 99% of accuracy only by using 3064 MRI images, i.e., the tactic is unable to handle vast amount of dataset for effective classification.

3 Limitations of Existing System

In the current arrangement of extraction of brain tumor, growth from CT images cancer part is disguised from the CT sweep of the mind. The proposed arrangement does likewise; advise the client about insights of cancer utilizing fundamental picture handling procedures.

The strategies incorporate commotion evacuation and honing of the picture alongside fundamental morphological capacities, disintegration and expansion, to get the foundation. Deduction of foundation and its negative from various arrangements of pictures brings about separated growth picture. The distinction in the proposed arrangement with existing arrangement is plotting form and c-mark of the cancer and limit which furnishes us with data identified with the growth that can help in a superior perception in diagnosis cases. This cycle helps in recognizing the size, shape and position of the cancer. It helps the clinical staff with welling the patient to appreciate the sincerity of the development with the help of different concealing stamping for different levels of tallness.

4 Problem Definition

This paper is identification of the brain tumor also the presence of a tumor through combining numerous technique to offer a foolproof technique of tumor detection in CT scan. The techniques applied are filtering, evaluation adjustment, negation of a photograph, photograph subtraction, erosion, dilation, threshold and outlining of the tumor. The cognizance of this challenge is CT mind pix' tumor extraction and its illustration in less complicated shape such that its miles comprehensible through everyone. Humans have a tendency to apprehend colored pix higher than black and white pix, thus, we are the use of colorings to make the illustration less complicated sufficient to be understood through the affected person in conjunction with the clinical team of workers. The goal of these paintings is to deliver a few beneficial facts in less complicated shape in the front of the users, specifically for the clinical team of workers treating the affected person.

This paper considers CT image and bunch of rules as an approach to achieve extricated photo of the growth. The resultant photo could fit for offer realities like size, estimation and capacity of the cancer, plotting form and c-mark of the growth, and its limit bears the cost of us with realistic related with the cancer which could show gainful for different cases with a reason to offer a higher base for the group of laborers to decide the relieving system. Plotting shapes f plot and c-mark plot of the disease and its breaking point will supply smooth data to the clinical gathering of workers on account of the truth people see pix higher with the assistance of various colorings for first class levels of power giving 3D insight from 2D photograph.

5 Proposed Work

In this paper, consider the algorithm is a set of image processing fundamental procedures [9, 10]. A set of noise-removal functions accompanied with morphological operations that result in clear image of tumor after passing through high-pass filter is the basic idea behind the proposed algorithm. The set of morphological operations used will decide the clarity and quality of the tumor image.

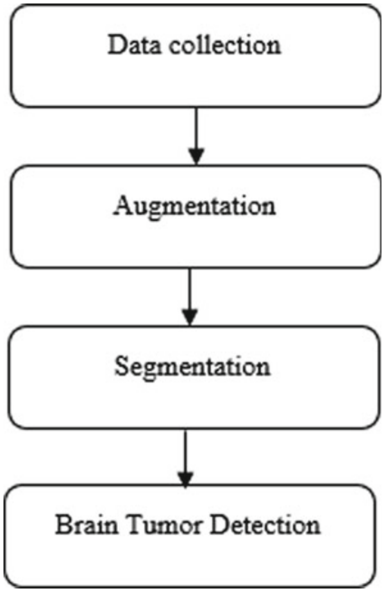
A GUI is created in the Python offering the proposed application of extracting the tumor from selected brain image and its visualization using contour plot. Without having to deal with the code, medical staff can select the CT image and study the extracted tumor along with its boundary from contour and c-label options. The GUI also contains options for zoom-in, zoom-out, data cursor for coordinates and prints the selected image.

Figure 1 shows the block diagram of proposed method for brain tumor detection includes.

5.1 Dataset

Dataset is available in the Google which considers dataset from kaggle or BraTS 2017 preparing dataset is utilized in the proposed strategy for mind growth division, which comprises of 210 HGG and 75 LGG cases, was explained physically by one to four raters and all divisions were supported by master raters. For each patient, a T1-weighted, a post-contrast T1-weighted, a T2-weighted and a FLAIR MRI were provided [11]. The MRI originates from 19 institutions and was acquired with different protocols, magnetic field strengths and MRI scanners. Each tumor

Fig. 1 Block diagram of the brain tumor detection



was segmented into edema, necrosis and non-enhancing tumor and active/enhancing tumor.

5.2 Augmentation

Subsequent to gathering the pictures, information increase measure is gone through like turn, level or vertical flipping and shear change. The organizations are significantly dependent on huge information to keep away from over fitting. In clinical picture examination area, tragically production of bigger datasets is significant test. To address this issue, information expansion strategy is done which works on the quantity of tests in the preparation dataset.

5.3 Segmentation

Here, we developed one new model name of that model is conventional neural network. In the wake of expanding the information, division is done by utilizing 3D-CNN. The 3D-CNN strategy pictures division which predicts every pixel's class. The 3D convolution is accomplished by convolving a 3D piece to the 3D shape framed by stacking numerous bordering outlines together and by this development, the element maps in the convolution layer are associated with various adjoining outlines in the past layer, so subsequently catching movement data. The element maps in the slope *x* and inclination *y* channels are gotten by processing angles along the level and vertical headings, individually or on every one of the seven info outlines, and the pick stream *x* and select stream *y* channels contain the optical stream fields, along the flat and vertical bearings, separately, figured from adjoining input outlines.

5.4 Brain Tumor Detection

This is the final step of above diagram, this step helps us to identify the size, shape and position of the tumor. It helps the medical staff as well as the patient to understand the seriousness of the tumor with the help of different color labeling for different levels of elevation. A GUI for the contour of tumor and its boundary can provide information to the medical staff on click of user choice buttons.

6 Result and Discussion

Finally, this paper focuses on result and discussion, and character of growth is a vital part in its treatment. But previously having so many loop poles to identify the brain tumor perfectly.

- Difficult to identify the size of the tumor.
- Trouble to recognize the shape of the tumor.
- Inconvenience to perceive the position of the tumor.
- Clinical staff hard to comprehend the earnestness of the patient.
- Patient and patient family members are likewise hard to comprehend the reality of the patient.
- There could be no legitimate diverse shading naming for various degrees of rise.

To minimize the above limitations, this paper introduces contour plot and c-label plot algorithm of tumor in CT images.

The following results showcase the outputs received after each step in the algorithm. Basic pre-processing transformations include:

- Changing the picture to grayscale, as we need to discover form of the last picture which deals with grayscale pictures.
- Applying low-pass filter, to remove any noise, if present, in the image.
- Applying high-pass filter, to obtain sharpened image with clear-defined boundaries.

Above diagrams show tumor having images and no tumor image, by using contour plot and c-label plot algorithm used to provide 3D visualization from the 2D image as shown in Fig. 2 with following advantages:

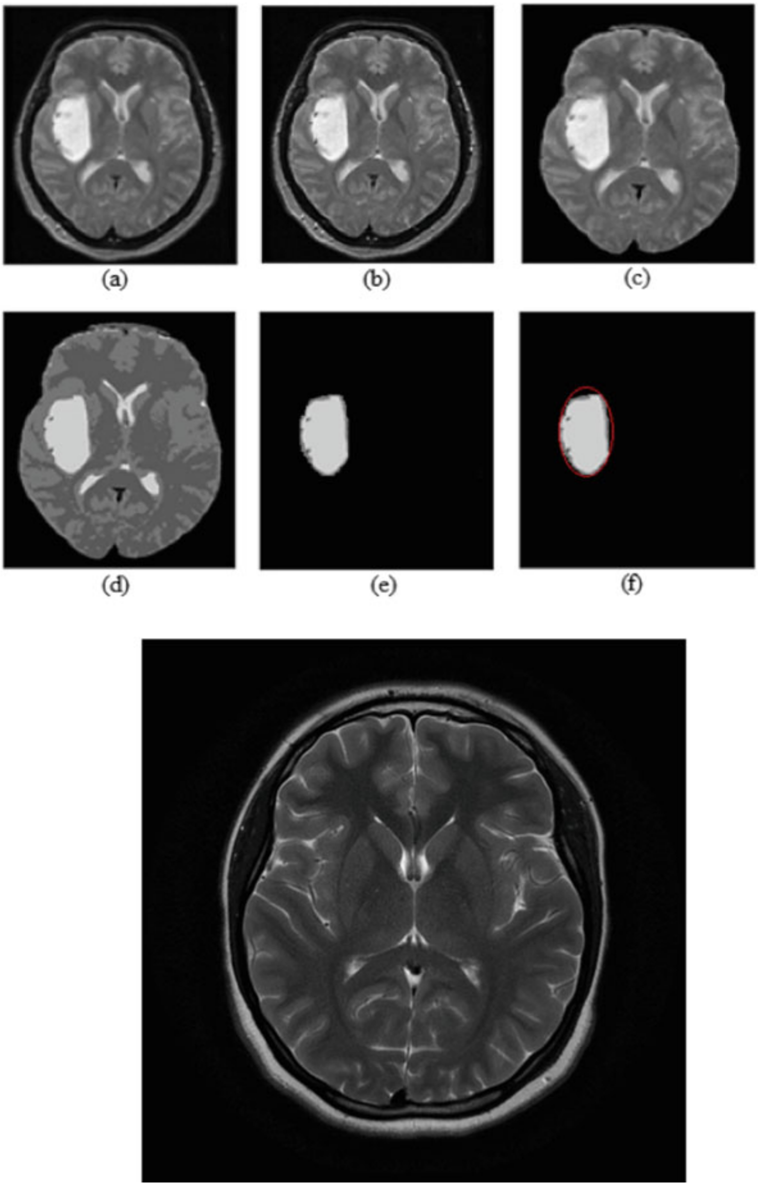


Fig. 2 Last image not having tumor, **a**–**f** images are having tumor recognize the size, shape and situation of the tumor

- Simple to recognize the size of the tumor.
- Easy to recognize the shape of the tumor.
- Comfort to see the situation of the tumor.
- Clinical staff easy to comprehend the reality of the patient.
- Patient and patient family members are likewise easy to comprehend the reality of the patient.
- There is unmistakable shading naming for various degrees of rise then, at that point straightforward trouble of the growth.

7 Conclusion and Future Enhancement

Analyzing brain tumor using CT scan is considered as the most significant concept for research and analysis. The proposed method is inputted with grayscale images of brain that contain tumor, the image is processed through various stages of morphological operations like filtering, contra adjustment, erosion, dilation, etc., through Python programming. Hence, tumor is outlined in the original image and clearly demarcated. Contour plot and c-label plot are created to provide 3D visualization from the 2D image and a GUI is also developed which enables the above application with a user-friendly interface, possible extension of the presented work could use more features. Also it would be beneficial to connect the system to cloud storage of patient's information in hospital. And the application can be extended to accessibility and usability through mobile phones.

Acknowledgements We would like to thank the GIET University and MRIET College for the immense support for our research work to implement.

References

1. Zhou Z, He Z, Jia Y (2020) AFPNet: A 3D fully convolution neural network with Atrous-convolution feature pyramid for brain tumor segmentation via MRI images. Neurocomputing 402:235–244
2. Javeria A, Muhammad S, Mudassar R, Mussarat Y (2018) Brain tumors detection using feature fusion and machine learning. J Amb Intell Human Comput. https://doi.org/10.1007/s1265\s2-018-1092-9
3. Mustaqeem A, Jived A, Fatima T (2012) A productive cerebrum cancer location calculation utilizing watershed and thresholding-based division. Int J 4
4. Narmatha C, Eljack SM, Tuka AARM, Manimurugan S, Mustafa M (2020) A hybrid fuzzy brain-storm optimization algorithm for the classification of brain tumor MRI images. J Amb Intell Human Comput:1–9
5. Rajeswari R, Ananda Kumar P (2011) Picture division and recognizable proof of cerebrum cancer utilizing FFT strategies of MRI pictures. ACEEE Int J Commun 02(02)
6. Zeineldin RA, Karar ME, Coburger J, Wirtz CR, Burgert O (2020) DeepSeg: deep neural network framework for automatic brain tumor segmentation using magnetic resonance FLAIR images. Int J Comput Assist Radiol Surg 15(6):909–920

7. Chen S, Ding C, Liu M (2019) Dual-force convolution neural networks for accurate brain tumor segmentation. Pattern Recogn 88:90–100
8. Ismael SAA, Mohammed A, Hefny H (2020) An enhanced deep learning approach for brain cancer MRI images classification using residual networks. Artif Intell Med 102:101779
9. Mohammad MH, Mohammad MH, Mohammad MH, Mohammad MH, Mo (2019) For brain tumors classification, a hybrid feature extraction technique with a regularized extreme learning machine is used. IEEE Access:36266–36273
10. Aboul EH (2019) Machine learning paradigms: theory and application. Springer Nature, Geneva
11. Ghosh D, Natrajan P, Sandeep KN (2013) Detection of tumor in mammogram images using broadened local minima threshold. Int J Eng Technol 5(3)

Empirical Study on Method-level Refactoring Using Machine Learning

Rasmita Panigrahi, Sanjay Kumar Kuanar, and Lov Kumar

Abstract Because of the importance of software refactoring for software code quality and stability, this research primarily emphasizes whether refactoring can be vital to identify probable software components for future refactoring. Modularity, reusability, modifiability, maintainability, and service-oriented development may all be improved with refactoring. This fact encourages academics to develop a new and improved machine learning paradigm for restructuring OO software. We have made a multi-purpose optimization effort to assess the OOP-based software systems or components refactoring in this work. This research intends to exploit and optimize OOP software metrics to examine code quality by performing refactoring. Our objective is to develop a highly resilient and efficient ensemble computing model for refactoring prediction at the method level into a machine learning framework using software metrics as features. The focus is on applying enhanced state-of-art data acquisition, data preprocessing, data imbalance resilient re-sampling, feature extraction, and selection, followed by improved ensemble-based classification. This work will also focus on the types of project work for different kinds of classification.

Keywords Method-level refactoring · Machine learning · Software metrics

1 Introduction

For scientific, social, and business communities, exponential growth in software computing technologies has opened up new opportunities to use them for improved productivity and procedural improvement in recent years. As a result, it has revitalized human life for making more efficient decisions to lead a better life. Thus, software systems have played a decisive role. Functionally, software systems can be defined

R. Panigrahi (✉) · S. K. Kuanar
School of Engineering and Technology, Department of Computer Science and Engineering (CSE), GIET University, Gunupur, Odisha 765022, India
e-mail: rasmita@giet.edu

L. Kumar
Department of Computer Science and Information Systems, BITS-Pilani, Hyderabad Campus, Room #H-134, Jawahar Nagar, Shameerpet Mandal, Hyderabad 500078, India

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_57

as one of the critical inventions made so far to execute specific intended tasks to serve varied applications, concerning everything from science and technology to the business and financial sectors to the healthcare industry everything else besides. Undeniably, software systems can be stated as an unsubstitutable need for human life that cannot be ignored due to irreplaceable significance to modern human society's needs. As a result of the enormous demand for current socio-scientific requirements, the software industry has expanded in scope, inviting enormously large-scale corporate communities to investigate improved technologies for better and enhanced performance. However, such demands are inevitably highly correlated to software design and reliability.

With the emergence of software computing innovations, ensuring the system's reliability has become a challenging task for which industries and allied software developers have been making vital efforts. Identifying software design, architecture, probable fault conditions, maintainability, bug probability, smells, replication, etc., can prevent a software system from becoming corrupted or failing during the software development life cycle (SDLC). In such a case ensuring Quality of Service (QoS) delivery with computationally efficient software development can be significant. To examine the efficacy of software systems assessing the entire system's refactoring can be of great significance as it can avoid any fault or design issues in the future, thus making the software system more efficient. When it comes to object-oriented software, refactoring is the process of making changes to a system's internal structure without affecting its functionality or performance. It means that the software system will be transformed into clean code and design. Software system quality and maintenance can be improved by implementing this technique. Refactoring can be done at the method, class, or package level, depending on the situation. After each change, the refactoring process performs step-by-step to ensure that everything continues to run smoothly and efficiently. Our source code metrics need to be computed based on class-level and method-level surveys to forecast the percentage of methods or classes to be refactored in the machine learning framework. Using software, refactoring can improve OOP's internal structure while still retaining its external behavior and functionality. Because code is constantly being modified to meet new requirements, the source code deviates from its original plan structure. Finding bugs in software code is difficult because the code becomes multifaceted, difficult to read and debug, and even too difficult to improve upon as a result of this. Software refactoring strategies let us remove odors that harm the program's operation while making it more comprehensible and extensible. The identification of code that requires refactoring is a significant challenge in the context of refactoring. It increases the difficulty of the task for the researchers.

2 Literature Review

Refactoring in software engineering literature is primarily discussed in this part as mentioned below. This section looks at the various approaches to quality measurement that are currently available in the aspect-oriented system. Earlier methods for evaluating the aspect system's external quality attributes were examined in this section. The following sections go into greater detail about the various efforts and outcomes that have been achieved. Eunjong et al. [1] accompanied a systematic review of the literature to detect software refactoring. The authors researched to make refactoring detection as straightforward as possible for researchers and practitioners. The primary aim of introducing refactoring methods, algorithms, and models is to make the code easy and clean. In conducting a systematic literature review, the researchers considered no hypotheses, and no data set was used in the study. Therefore, multiple techniques can be combined and evaluated quantitatively using the evaluation process proposed by the authors. Alexander et al. [2] concentrated their efforts on figuring out how the system's internal quality attributes (cohesion, coupling, complexity, inheritance, and size) were affected by the source metric. The authors primarily used refactoring techniques, i.e., move method, inline method, move field, extract class, and in their work. In addition, 23 Java projects (open-source) from GitHub with 29,303 refactoring operations were provided to the authors as part of the experiment, which they had to complete during the investigation. Because of this, the authors made a hypothesis: the refactors internal properties will impact it.

Dallal et al. [3] As the first step in this research, they conducted a systematic literature review to identify and analyze existing findings on the impact of refactoring on software quality and ensure that the findings are consistent with each other. According to the authors, the literature on bed smell code was primarily dominated by empirical study articles (object-oriented code refactoring). When compiling and analyzing their findings for their study on the impact of refactoring on software quality, the authors used vote counting. The authors' conclusions are as follows:

(1) Refactoring tools such as the MOVE method and the EXTRACT method are used in most refactoring scenarios.
(2) The majority of refactoring tools are based on a specific quality characteristic (complexity, cohesion, size, inheritance).
(3) Software quality indicators for indicating software quality are studied using statistical techniques in a small number of studies.
(4) The articles related to empirical study are composed of data sets originating from various locations and created using a variety of programming languages of varying sizes.

Kodhai et al. [4, 5] proposed a new model (clone manager) to detect the clones and matched its results with those of three other open-source java tools (Aries, Refactoring Crawler, and Cedar) (Aries, Refactoring Crawler, and Cedar). Additionally, they provided a conceptual diagram of the model and an algorithm for three refactoring tools that are currently available. Finally, they discovered the answer by counting the number of times the methods (pull-up, extract, and move) appeared in different integrated development environments (IDEs). Kaur et al. [6] A conceptual model diagram and algorithms for three currently available refactoring tools were provided. They have made the software projects available for download because of the data set collected from GitHub repository. It has been argued that low-maintenance code, thorough reworking, and the influence of refactoring on the software maintainability index can all be combined. Vidal et al. [7] A method known as the brain method was developed to consolidate intelligence about classes. This method stands out for being long and complex, making it difficult for programmers to understand.

Nevertheless, this is a recommended system that suggests refactoring options and enables the elimination of unpleasant odors by applying a search-based approach based on simulated annealing. More than 60% of the errors were automatically fixed when using open-source Java application projects as a benchmark. Thirty five industrial programmers were polled, and they all agreed that Bandago's refactoring proposal outperformed the third-party refactoring tool in terms of usefulness.

Bavota et al. [8] A method book and a new approach have been developed to eliminate the odors associated with feature envy in code writing. They used semantic and structural dimensions to determine the degree of similarity between various methods as their primary focus. In addition, the degree of cohesiveness, coupling, and design length is often utilized to judge whether or not a strategy can be extracted.

Liu et al. [9] also discussed conflict-aware scheduling schemes and metrics-based scheduling schemes used during software project development and proposed refactoring rules, referred to as heuristic rules. The multi-objective optimization (MOO) model is what used to measure this metric. Three-step refactoring was discussed by AnastasiosTsimakis and colleagues [10]. They created a tool that can detect the refactoring detection pattern and alert the user. This tool can be used to identify the eleven different types of refactoring patterns. The process of refactoring can be accomplished through a variety of operations. The refactoring function is the most popular among authors out of the various extract methods.

Silva et al. [11] looked at some Java projects on GitHub for refactoring techniques and reasons given by the developers for refactoring their code. They should then generate a list of 44 distinct motives for each of the 12 refactoring strategies. They discovered that refactoring works better when requirements are changed than when code smells are present. The authors concentrated on strategies for refactoring extract methods because it is a versatile procedure. Finally, the author concluded that refactoring tools depend on the developers' interactive development environment.

Charalampidou et al. [12] highlighted the benefits of the extract method and showed lengthy methods (smelly method) utilized for refactoring. The authors present a methodology (supported by an apparatus) to extract source code fragments that operate together to offer specified functionality; the authors present a method (supported by an appliance). A contemporary and open-source environment have been empirically validated. The open-source data (about 500 LoC per) received a 93% evaluation rate. Wen-mei LIU et al. [13] presented the extract approach for refactoring. They conducted some research and approved or disapproved of the extract technique's critical application in refactoring. Initially, the authors interviewed 25 software developers. There was a 28% primary motive for extract method refactoring and a 16% to resolve the clones. They have got the conclusion that identifying needs early in the software development process is critical. In this paper, Xu et al. [10], the innovative framework that learns refactoring method criteria from open-source software analyzed and evaluated the three primary techniques available today: SEMI, J-Extract, and Deodorant. They discovered that their strategy was statistically effective. For the only extract technique, they presented GEMS (generate-and-rank recommender). It uses a probabilistic approach to count the scoring function.

Pandian et al. [14] discussed the significance of AI and its presentation in an intelligent warehousing atmosphere. Kumar et al. [15] explored the possibility of building tool support to assist software engineers in identifying methods and classes that need refactoring. They have employed the least-squares support vector machines (LSSVM) to achieve their results. Principal component analysis (PCA) as a feature extraction technique. When dealing with unbalanced data, synthetic minority over-sampling technique (SMOTE) comes in handy. They have gathered information from the tera-PROMISE. The study's primary objective was to determine the consequence and association effect of SMOTE and PCA on the performance of LSSVM kernels. They discovered that the value of AUC was 0.96.

In the paper, Vedurada et al. [16], the control field is primarily concerned with the behavior of the classes, as the name suggests. Their primary focus was on the CF algorithm, which was used to identify and crop them. They enlightened that they would primarily replace classes with subclasses and states. Panigrahi et al. [17] implemented a machine learning algorithm for refactoring methods at the method level. To estimate accuracy and AUC, they used variants of Naive Bayes classifiers such as multinomial, Bernoulli, and Gaussian classifiers, among others. They have collected the data set from the tera-PROMISE data storage facility. This study by Kumar et al. [18] aims to improve refactoring approaches. A machine learning approach is used to anticipate the need for refactoring based on 25 source code metrics. The authors also provide a method of testing the proposed system's performance using data from the tera-PROMISE repository. This work suggests ten machine learning classifiers and three sampling approaches cope with imbalanced data. For each sample strategy and classifier, the author used box plots to compare the mean accuracy. On the class level, two clustering approaches (clustering with a variable no. of classes and fixed number of classes) and one graph transformation approach are reviewed by Ismail [19]. The author used three different clustering approaches [SLINK, CLINK, and WPGMA] and one adaptive K-ANN. This is because the A-KNN technique reduces coupling between packages while boosting package cohesiveness. An approach's objectivity, level of refactoring, and supportability are evaluated against the approach's complexity and validity. Vijayakumar et al. [20] completed the systematic investigation of capsule neural networks in a variety of applications. Liu et al. [21] have been experimentally demonstrated that manually identifying code smell is a time-consuming task, and tools basing up on deep learning for detecting code smells have been developed. First, the user must choose the metrics used for prediction purposes, which can be difficult. Even if we decide, determining which one will be the best is a difficult task once again to accomplish. They proposed a DL-based approach that is simple to use and can easily detect code smells. There is no need for human intervention in developing this tool, which automatically detects long methods, large classes, and feature envy.

3 Literature Gap Analysis and Aim of Research

The above discussion reveals that refactoring analysis and detection systems can be of paramount significance; however, automated AI-based strategy development has been a hurdle for academic industries. Though a few efforts have been made to perform code smell detection, maintainability detection, etc., its generalization as a global solution remains questionable. Though not many efforts have been made toward exploiting structural features to detect refactoring probability and malicious entry identification in software systems, identifying the most significant structural constructs of features has motivated scientific society. Object-oriented programming (OOP) metrics are standard in OOP-based software systems. Still, minimal effort has been made to segment the most powerful features of the software component that can have a strong or significant impact on the refactoring classification or accuracy of refactoring probability identification.

On the contrary, not all structural features can be crucial to perform refactoring prediction or analysis. It motivates the research community to explore and exploit different possible structural features such as OOP metrics for their respective significance toward refactoring prediction or analysis. Though a few efforts have been made to use basic machine learning methods for refactoring analysis, their efficacy as a globally acceptable solution remains suspicious. The reason behind such an approach is the difference in classification accuracy. For example, logistic regression, K-NN, SVM, and ANN can have different accuracy even with the same data (OOP metrics features). Since this as motivation, ensemble computing methods have been recommended; however, the selection of suitable base classifiers has always been a challenge. It motivates authors to explore and achieve an optimal choice of base classifiers and ensemble design. Unlike conventional ensemble models such as boosting, AdaBoost, and random forest identifying optimal classification models can be vital toward refactoring analysis and classification. Cross-project refactoring prediction is a well-known research topic that entails developing a model from a project database and then testing it on a database from a different project, among other things. Specifically, machine learning and deep learning techniques are used in conjunction with the proposed attribute selection phases to predict refactoring at the method level. Projects have long been recognized as virtual repositories of organizational knowledge and innovation, particularly in the public sector. It is the creation and transfer of knowledge that takes place within a project that is known as intra-project learning. Knowledge transfer is required in project-based organizations (PBOs) for inter-project and intra-project activities to achieve success. Summarily, the following key factors can be the driving force or motivation behind the current study.

1. Identification of key structural qualities (such OOP metrics) for refactoring classification.
2. Recognizing most efficient features for refactoring analysis and classification.
3. Recognizing most optimal and robust ensemble model for refactoring classification or prediction.
4. Recognizing the cross-project model for refactoring classification or prediction.
5. Recognizing the intra-project model for refactoring classification or prediction.

4 Research Methodology

Considering the significance of software refactoring for software code quality and reliability, this research primarily emphasizes assessing whether refactoring can be vital to identify probable software components for getting refactored in the future. Refactoring can help software enrich modularity, reusability, modifiability, maintainability, and service-oriented development motivates the researcher to design a novel and enhanced machine learning paradigm for refactoring the object-oriented software program. With this motivation, we have made a multi-purpose optimization effort to assess refactoring of the OOP-based software systems or components in this research work. This research intends to exploit and optimize OOP software metrics to examine code quality by performing refactoring, where the focus is made on applying enhanced state-of-art data acquisition, data pre-processing, data imbalance resilient re-sampling, feature extraction, and selection followed by enhanced ensemble-based classification.

1. **Issue 1**: Machine learning algorithms commonly fail or provide unjustified findings when used to classification data sets with an uneven class distribution. Data sampling is a set of techniques for balancing or improving the distribution of classes in a learning environment. Standard machine learning methods can be trained directly on the balanced data set without extra processing. This enables a data preparation strategy to successfully solve the difficulty of unbalanced classification, even when class distributions are extremely skewed. Various data sampling strategies exist, each with its own benefits and drawbacks. The optimal technique to solve all classification problems and apply all classification models, like choosing a prediction model, is different for everyone.

Solution: To implement the most efficient technique for handling data imbalance for dimensionality reduction in high dimensional data. We plan to use several data sampling techniques to solve the data imbalance issue for our proposed model. The data sampling techniques that can be used are:

- SMOTE
- LSSVM

- BLSMOTE
- SVSMOTE.

2. **Issue 2**: When it comes to machine learning and predictive modeling, feature selection is extremely important. The dimensionality reduction strategy includes a variety of measures, one of which is dimensionality reduction. Selection of relevant characteristics for processing without any transformation is what feature selection is primarily characterized as. It is also referred to as attribute selection or variable selection in some circles. It aids in the selection of the most appropriate features from among the many options available in RQ2 of Fig. 1.

The selection of features can be made either manually or automatically. Because some features are expensive to obtain, narrowing down the list of available features is beneficial. When features are transformed, the measurement units associated with them are lost. However, when it comes to feature selection, measurement units are maintained. As a result, the model's accuracy improves as a result of it. It also cuts down on time it takes for the model to become proficient. After that, it deletes everything that is not necessary.

Solution: It is necessary to identify the most acceptable and statistically significant OOP software metrics for method-wise refactoring prediction in large-scale OOP software systems, as well as their statistical significance. The Wilcoxon rank-sum test, significant predictor test, correlation test, principal component analysis test, and

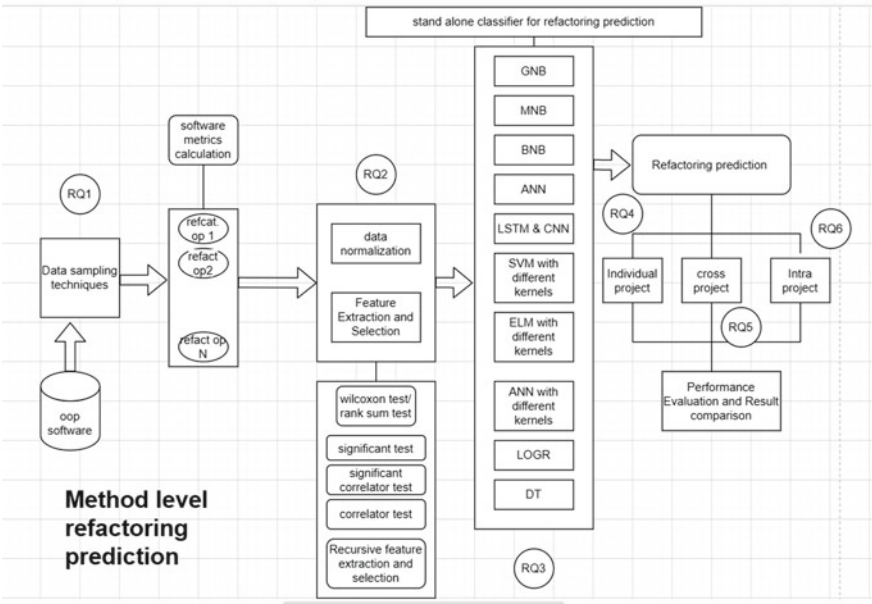


Fig. 1 Proposed model for refactoring prediction

other recursive feature extraction processes for large-scale object-oriented software systems are examples of such tests.

3. **Issue 3**: Machine learning algorithms use various techniques to deal with large amounts of complex data to make informed decisions. Creating a model begins with training the machine learning algorithm on a training data set, and then used to generate results. Predictions and results are checked for accuracy before being presented to the audience. Then, the algorithm is re-trained over and over again until it produces the desired output. If the prediction does not turn out as expected, the process is repeated. This enables the machine learning algorithm to learn independently and produce an optimal answer that will gradually improve accuracy over time due to the learning process. The machine learning algorithm is activated once the desired level of accuracy has been achieved using the training data. In machine learning, ensemble methods combine several base models to produce a single optimal predictive model. Ensemble approaches integrate many machine learning techniques to reduce variance (bagging), increase bias (boosting), and improve predictions (stacking). Deep learning is a machine learning method based on artificial neural networks and representation learning. In machine learning, it uses artificial neural networks and representation learning.

Solution: Refactoring prediction models for big OOP software systems should be implemented using machine learning. We are proposed to use optimized machine learning classification algorithms and ensemble techniques for better prediction in terms of accuracy, precision, recall, AUC, and *F*-measure. The following machine learning algorithms can be implemented for refactoring prediction, and we have already implemented some of the algorithms like Naïve Bays, LSTM, and CNN. The rest of the algorithms we are going to use in our future work.

- Logistic Regression
- Naïve Bayes
- Stochastic gradient descent
- K-nearest neighbors
- Decision tree
- Random forest
- Support vector machine
- ELM based
- LSTM and CNN.

4. **Issue 4**: If there is enough data to train the model, software defect prediction is effective for projects. However, innovative software projects and other businesses do not face this issue. As long as we are using a single model for training and testing, it is critical to transfer predictions between models without losing any information. Cross-project refactoring is the term for this process. CP refactoring transfer learning prediction performs a large-scale defect prediction. Unfortunately, cross-project prediction becomes more difficult for real-world applications.

Solution: We aim to implement a machine learning-based refactoring prediction model for large-scale OOP software systems with a cross-project learning mechanism. We are proposed to use optimized machine learning classification algorithms and ensemble techniques for better prediction in terms of accuracy, precision, recall, AUC, and *F*-measure.

Issue 5: Only a few studies have looked into the knowledge transfer activities within and between projects. On the one hand, face-to-face interactions in projects can transfer knowledge, while IT-based interactions can transfer knowledge. On the other hand, there are tools. Even though companies have allocated a significant amount of resources to IT tools, it has been discovered that they are not always effectively utilized, and people prefer to look for alternative solutions. Therefore, face-to-face social interactions are used to gain knowledge.

Solution: Our future aim is to develop an OOP software refactoring prediction model based on machine learning for large projects with the intra-project scheme. We are proposed to use optimized machine learning classification algorithms and ensemble techniques for better prediction in terms of accuracy, precision, recall, AUC, and *F*-measure.

5 Conclusion

We make use of source meter for Java's source code metrics. Machine learning algorithms make use of the source code metrics as features or independent variables. The source code measurements are fed into the machine learning algorithms, which produce a binary class (refactoring or not). We have already conducted Wilcoxon rank-sum test in some of our published papers. In the future, we will conduct a correlator test, significant Correlated test, and other tests during the pre-processing stage for feature extraction and selection purposes. We identify valuable metrics for refactoring prediction using metrics selection. We find a statistically significant difference in performance between some classifiers and imbalance learning strategies. Therefore, it is feasible to forecast method-level refactoring using source code metrics and machine learning classifiers in individual, cross-project, and intra-project. We aim to build a highly robust and efficient ensemble computing model for refactoring prediction at the method level with software metrics as the machine learning framework features.

References

1. Choi E, Fujiwara K, Yoshida N, Hayashi S (2018) A survey of refactoring detection techniques based on change history analysis. arXiv:1808.02320

2. Chávez A, Ferreira I, Fernandes E, Cedrim D, Garcia A (2017) How does refactoring affect internal quality attributes? A multi-project study. In: Proceedings of the 31st Brazilian symposium on software engineering. ACM, pp 74–83
3. Al Dallal J, Abdin A (2017) Empirical evaluation of the impact of object-oriented code refactoring on quality attributes: a systematic literature review. IEEE Trans Softw Eng 44(1):44–69
4. Kodhai E, Kanmani S, Kamatchi A, Radhika R, Saranya BV (2010) CloneManager: a tool for detection of type1 and type2 code clones. In: International conference on business administration and information processing. Springer, Berlin, Heidelberg, pp568–570
5. Kodhai E, Kanmani S (2014) Method-level code clone modification environment using CloneManager. In: Modern trends and techniques in computer science. Springer, Cham, pp 529–539
6. Kaur G, Singh B (2017) Improving the quality of software by refactoring. In: 2017 international conference on intelligent computing and control systems (ICICCS). IEEE, pp 185–191
7. Vidal S, Zulliani S, Marcos C, Pace J (2018) Assessing the refactoring of brain methods. ACM Trans Softw Eng Methodol (TOSEM) 27(1):2
8. Bavota G, Oliveto R, Gethers M, Poshyvanyk D, De Lucia A (2013) Method book: recommending move method refactorings via relational topic models. IEEE Trans Softw Eng 40(7):671–694
9. Liu H, Li G, Ma ZY, Shao WZ (2008) Conflict-aware schedule of software refactoring. IET Softw 2(5):446–460. https://doi.org/10.1049/iet-sen:20070033
10. Tsimakis A, Zarras AV, Vassiliadis P (2019) The three-step refactoring detector pattern. In: 24th European conference on pattern languages of programs (EuroPLoP'19) p 9, July 3–7, 2019, Irsee, Germany. ACM, New York
11. Silva D, Tsantalis N, Valente MT (2016) Why we refactor? Confessions of GitHub contributors. In: Proceedings of the 2016 24th ACM SIGSOFT international symposium on foundations of software engineering. ACM, pp 858–870
12. Charalampidou S, Ampatzoglou A, Chatzigeorgiou A, Gkortzis A, Avgeriou P (2016) Identifying extract method refactoring opportunities based on functional relevance. IEEE Trans Softw Eng 43(10):954–974
13. Liu W, Liu H (2016) Major motivations for extract method refactoring: an analysis based on interviews and change histories. Front Comp Sci 10(4):644–656
14. Pandian AP (2019) Artificial intelligence application in a smart warehousing environment for automated logistics. J Artif Intell 1(02):63–72
15. Kumar L, Sureka A (2017) Application of LSSVM and SMOTE on seven open-source projects for predicting refactoring at the class level. In: The 2017 24th Asia-Pacific software engineering conference (APSEC). IEEE, pp 90–99
16. Vedurada J, Krishna Nandivada V (2018) Identifying refactoring opportunities for replacing type code with subclass and state. Proc ACM Progr Lang 2(138):28. https://doi.org/10.1145/3276508
17. Panigrahi R, Kuanar SK, Kumar L (2020) Application of Naïve Bayes classifiers for refactoring prediction at the method level. In: 2020 international conference on computer science, engineering and applications (ICCSEA). IEEE, pp 1–6
18. Kumar L, Satapathy SM, Murthy LB (2019) Method level refactoring prediction on five open source java projects using machine learning techniques. In: Proceedings of the 12th innovations on software engineering conference (formerly known as India software engineering conference). ACM, p 7
19. Keshta IM (2017) Software refactoring approaches a survey. Int J Adv Comput Sci Appl 8(11):542–547
20. Vijayakumar T (2019) Comparative study of capsule neural network in various applications. J Artif Intell 1(01):19–27
21. Liu H, Jin J, Xu Z, Zou Y, Bu Y, Zhang L (2019) Deep learning-based code smell detection. IEEE Trans Softw Eng. https://doi.org/10.1109/TSE.2019.2936376, http://www.ieee.org/publications_standards/publications/rights/index.html

State of the Art of Ensemble Learning Approach for Crop Prediction

Shraban Kumar Apat, Jyotirmaya Mishra, K. Srujan Raju, and Neelamadhab Padhy

Abstract Agriculture has a captious part in maintaining a large population. It plays a decisive role to forward our country's economic development. Crop cultivation has been the most prominent problem in recent days due to changes in weather patterns. This has a significant impact on crop productivity, either directly or indirectly. As a result, new technologies might be brought up for use in order to overcome this problem and uplift crop output. In this research proposal, we have explored IoT by choosing the smart farm and digital technology and explaining the management of heterogeneous data for agriculture. We proposed an IoT-HELE-based smart farming prediction and intelligent agriculture analytics model and a decision support system that effectively predicts crop production by utilizing cutting-edge machine learning and deep learning techniques. In this model, ensemble voting results in a more efficient, sustainable, and profitable agriculture enterprise. The multi-source dataset from the National Research Council (CNR), an ISTAT, and an IoT sensor will be analyzed. This work is presented through a new innovative idea after a rigorous literature review; presumably, it is valuable and increases the productivity of an agricultural firm.

Keywords IoT · Dataset · Ensemble machine learning · Deep learning

S. K. Apat (✉) · J. Mishra · N. Padhy
Department of CSE, GIET University, Gunupur, Odisha, India
e-mail: shraban.apat@giet.edu

J. Mishra
e-mail: jyoti@giet.edu

N. Padhy
e-mail: dr.neelamadhab@giet.edu

K. Srujan Raju
CMR Technical Campus, JNTUH, Hyderabad, India

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_58

1 Introduction

Everything in today's society is connected via the Internet, cloud, and IoT, which are used to implement innovative technologies in the social and productive sectors. We may use modern technology to handle numerous challenges, such as trash reduction, crop prediction, and production improvement, suggestions, and soil improvement, by obtaining and analyzing varieties of data.

We have proposed to use the three separate datasets in this study: a scientific dataset from the National Research Council (CNR), an ISTAT dataset, and an IoT sensors dataset.

For improved data administration and analysis in industrial firms, data from many sources are required, for which the authors employed various techniques.

(1) Machine-generated (MG) data: Information gathered via sensors, drones, and GPS.
(2) Process-mediated (PM) data: Information derived from business processes that refer to corporate entities.
(3) Human-sourced (HS): Data previously only available in books is now virtually completely digitalized.

From the historical data, extract the valuable information which gives better design and performance, for this, author used different tasks. We will use the dataset from the repositories mentioned above (Fig. 1) for effective crop prediction using various machine learning algorithms.

Task 1: The first task is to forecast future data (ISTAT dataset).

Task 2: Machine learning algorithms for missing data comparison (CNR).

Task 3: Using NN, linear, and polynomial regression models to reconstruct missing data from monitoring stations (IoT sensor dataset).

Task 4: Using decision tree and KNN models to reconstruct missing data (IoT sensor dataset).

Task 5: Sensor values are used to detect the defect monitoring stations (IoT sensor dataset).

2 Literature Survey

Khaki et al. [1] used DL with CNNs and RNNs for prediction of crop yield based on environmental parameters. This CNN-RNN model, along with RF, DFNN, and LASSO predicted corn and soybean yield 0. Using historical data from 2016 to 2018 over the whole Corn Belt of thirteen states in the United States. As average yields, this model had an RMSE of 9% and an RMSE of 8%. Yield performance, management, weather, and soil were the four kinds of data examined. This strategy could be used to solve other research challenges, such as predicting other crops using a hybrid system. Agarwal et al. [2] SVM is a machine learning algorithm, whereas LSTM and RNN

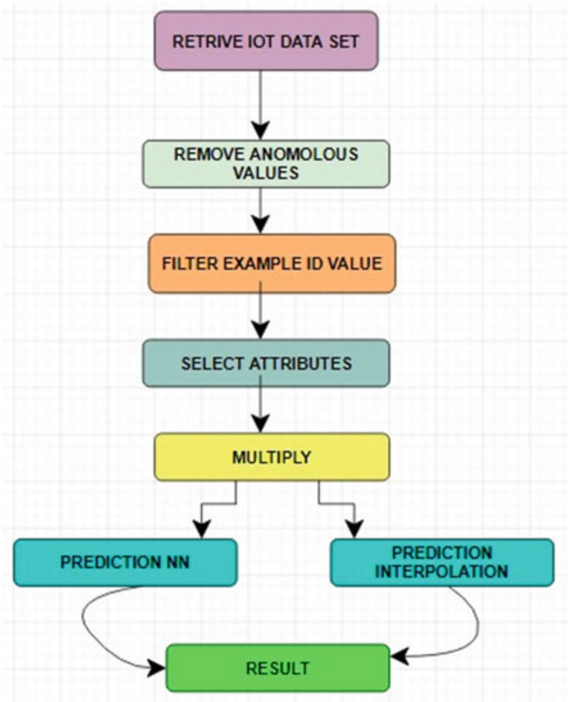


Fig. 1 IoT enabled data collection and crop/yield prediction mechanism

are deep learning algorithms. A Python-based system took input as climatic and soil conditions of land to predict a proper yield. About 741 kB of datasets is collected from kaggle. Prediction parameters used are temp—rainfall, pH, humidity, and area. Selected crops are pea, pigeon pea wheat, sugarcane, green gram rice, maize, millet, sugarcane, green gram, etc. Earlier applying ANN and RF algorithms, the accuracy is calculated as 93%. Whereas, in a present model using LSTM, RNN, and SVM, the accuracy is 97%.

Wang et al. [3] Meteorological and remote sensing data inputs were used to test the model on LSTM networks. CNN was used to model static soil properties in another division. The model was trained using detrended statistics yield data acquired between 1982 and 2015, and it was then evaluated after a year. Results show good performance with R^2 + RMSE of 0.76 and 721 kg/ha. The goal is to analyze the feasibility of DL-based yield prediction by long-term historical yield data. Alibabaei et al. [4] used DL, LSTM, and GR units with their extensions, called BLSTM and BGRU, to predict yields during end-of-the season. The expt. was tested for vegetables like potatoes and tomato in a geographical location of Portugal. The nonlinear interaction between irrigation, climatic data, and soil is implemented in this model. With an MSE of 0.016–0.038, the water content and expected yield. The BLSTM shows better performance than the other models with an R^2 score ranges 0.97–0.99. The problem is the time of training of the model BLSTM is higher. Gong et al. [5] suggested a TCN and RNN-based greenhouse crop yield prediction system.

For the growth of vegetables such as tomatoes, multiple datasets acquired from multiple places have been evaluated. The RMSEs between the predicted and actual crop yields show that the proposed approach achieves better performance than both traditional ML and classical DNN. Here, DNN-based methodology is proposed. The factors used are carbon dioxide, temp., humidity, radiation, etc. Datasets are collected from the UK, Newcastle agrowing site.

Khaki et al. [6] For the projected weather data, we used a DNN technique with an RMSE of 12% of the average yield and 50% of the standard deviation. With good weather data, the RMSE would be decreased to 11% of the average yield and 46% of the standard deviation. This model outperforms more prominent LASSO, SNN, and RT approaches, according to computational results. The datasets consist of 2268 exploratory hybrids planted in over 2247 locs between 2008 and 2016 across USA and Canada. Eight soil variables and 72 meteorological variables were used in the dataset. Bi et al. [7] Proposed model used GD, GA, NN, and GA-assisted DL solution method is applied for the crops prediction in two phases. Proposed method reduces the RMSE by about 10%. Shahhosseini et al. [8] proposed a model, i.e., crop modeling + ML which is a hybrid one for better predictions for corn yield prediction applied. Five ML techniques (LR, LASSO, Light GBM, RF, and XGBoost) in line with six ensemble models. Yield prediction RMSE can be reduced from 7 to 20% upon applying simulation crop model variables (APSIM) to ML models as an input feature. Author concluded that to improve more hydrological inputs are more essential and weather information alone are not sufficient. Chu et al. [9] BBI model is a new end-to-end prediction model for summer and winter rice yields that merges two BPNNs with an autonomous RNN. The model's RMSE for summer rice prediction (0.0044 and 0.0057) and winter rice prediction (0.0074 and 0.0192) are shown in the experimental data, with the layer in the network set to 06. Gopal et al. [10] The intrinsic relationship between MLR and ANN is investigated for effective rice crop prediction, and a hybrid MLR-ANN model is developed backpropagation training algorithm for feed forward artificial neural N/W. Both hybrid MLR-ANN and standard ANN computation times were calculated. The results suggest that the proposed hybrid MLR-ANN model outperforms traditional models in terms of accuracy.

2.1 Research Question

RQ#1: Can it be possible to design an efficient and robust machine learning model assisted by IoT for smart farming and related intelligent agriculture decisions?

RQ#2: Can it be used to create a strong heterogeneous ensemble learning environment (HELE) by leveraging the efficacy of various standalone machine learning algorithms.

RQ#3: Can the maximum voting ensemble (MVE) and base trained ensemble (BTE) paradigms be used to create a new HELE model?

RQ#4: Is it possible to create reliable heterogeneous ELE (HELE) machine learning-based HELE-MVE and HELE-BTE models for smart farming prediction, decision, and associated analytics systems?

RQ#5: Is it possible to adopt a multi-phase feature selection model to improve the suggested IoT-HELE model for smart farming prediction, decision, and allied analytics systems' accuracy, dependability, scalability, and computing efficacy?

RQ#6: Can it be possible to design the ensemble learning algorithms to predict crop and yield?

RQ#7: Is it possible to design the IoT-assisted model, which will use majority voting for crop and yield prediction for smart agriculture?

RQ#8: Can it be possible to use the deep learning technique to predict crop and yield prediction? If so, then which performance metrics will be effective for measuring crop and yield prediction.

2.2 Research Significance

Some of the key research significances hypothesized in this research is given as follows:

- Unlike classical efforts, in this research, heterogeneous datasets have been taken into consideration, which can provide better insight and multiple decision variables to assist optimal smart farming prediction, decision, and allied analytics systems. In significant existing efforts, a limited dataset has been used, while in this research, multiple IoT data collected from the benchmark and publically available sources have been considered. It can make overall analytics more productive and constructive to make a decision toward optimal smart farming and/or intelligent agriculture.
- To enhance overall performance in this research, efforts have been made to amplify data preprocessing, data sampling (in case of possible data imbalance condition), enhanced (multiple phased) feature extraction, and selection can play a decisive role in strengthening overall computational efficacy and reliability of the outcome.

3 Flowchart Representation for IoT Enabled Data Collection and Crop/yield Prediction

The IoT dataset is applied first, then inaccurate and missing values are eliminated, filters were used to locate monitoring stations (Fig. 1), characteristics are selected and combined, and lastly, the two machine learning sub-process blocks are executed, and the results are returned. To recover missing data from monitoring stations for the IoT sensor dataset, decision trees, polynomial models, and KNN models were utilized. This problem is a continuation of the previous task, except that it employs additional machine learning techniques while maintaining all of the prior task's hypotheses. Using an IoT sensor dataset, identify defect monitoring stations based on sensor values. The primary purpose of this work is to detect anomalies as fault stations by the use of threshold values to cluster monitoring stations geographically by frozen area, amplitude, and zinc attributes. Detect hardware faults with the use of IoT datasets and group the entire dataset.

4 Roadmap for Crop and Yield Prediction

Rapid Miner Studio is a visual workflow design tool, will be used throughout this exploration. Rapid Miner Studio is built in Java, and it will be used to manage large amount of data from the preprocessing phases through the ML methods applied to the data from the various sources (heterogeneous).

This software solution features a robust and user-friendly interface that enables rapid replication, reuse, and customization of workflows (Table 1).

5 Proposed IoT-HELE-based Smart Farming Prediction and Intelligent Agriculture Analytics Model

Figure 2 is called the smart farming prediction and intelligent agriculture analytics model. This model is called a fivefold model. The purpose of the suggested model is useful for decision-making tools for predicting crop yields and deciding which crops to plant and what precaution measure to be taken during growth of the crop in the specific season. Several machine learning approaches have been employed to improve agricultural yield prediction research. Convolution neural networks (CNNs) are the most commonly used deep learning approach, according to this supplemental data, with LSTM and deep neural networks following closely behind (DNN).

Phase #1: We have collected the benchmark datasets from different sources like CNR, ISTAT, KAGGLE, etc., (pertaining to smart farming, crop-yield production, and plant assessment). Once the dataset was collected, we need to follow the data preprocessing steps where significant test-(Wilkinson test), min–max normalization outlier analysis will be performed. The most important is to identify the relevant features from the heterogeneous dataset. Apart from these, we need to apply the algorithms (rank sum test, principle component analysis, Pearson correlation analysis, evolutionary computing-assisted feature selection).

Phase#2: Once the dataset is collected, need to follow the data preprocessing steps where significant test-(Wilkinson test), min–max normalization outlier analysis will be carried out. The most important is to identify the relevant features from

Table 1 Overall layout of crop and yield prediction

IoT-HELE machine learning-assisted prediction model for smart farming and intelligent agriculture	
Datasets	Agriculture datasets containing CNR, ISTAT, KAGGLE, etc., (pertaining to smart farming, crop-yield production, and plant assessment)
Preprocessing	Significant test-(Wilkinson test) Min–max normalization Outlier analysis
Data sampling (if required)	Random sampling Up-sampling Down-sampling
Feature selection	Rank sum test Principle component analysis Pearson correlation analysis Evolutionary computing-assisted feature selection
Machine learning	• KNN • K-means • Expectation–maximization (EM) • J48 pruned tree • RefTree • Decision trees (DT) • LR • LOGR • NB • SVM-Lin • SVM-Poly • SVM-RBF • LS-SVM • LS-SVM-Lin • LS-SVM-Poly • LS-SVM-RBF • Multivariate-adaptive-regression-spline (MARS) • ELM-Lin • ELM-Poly • ELM-RBF • ANN-GD • ANN-GDX • ANN-RBF • ANN-LM **Ensemble** • HELE-MVE • HELE-BTE
Deep learning techniques	• Long-short-term-memory (LSTM) • Recurrent-neural-network (RNN) • Deep neural network (DNN) • DBN • CNN • DAE
Output	• Predicted output/category/,etc., (based on data and intend)

(continued)

Table 1 (continued)

IoT-HELE machine learning-assisted prediction model for smart farming and intelligent agriculture	
Performance parameters	• Thus, the overall proposed model can be of considerable significance to achieve optimal smart farming and/or intelligent agriculture decisions • Accuracy, precision, F-measure, $F1$-score, etc.

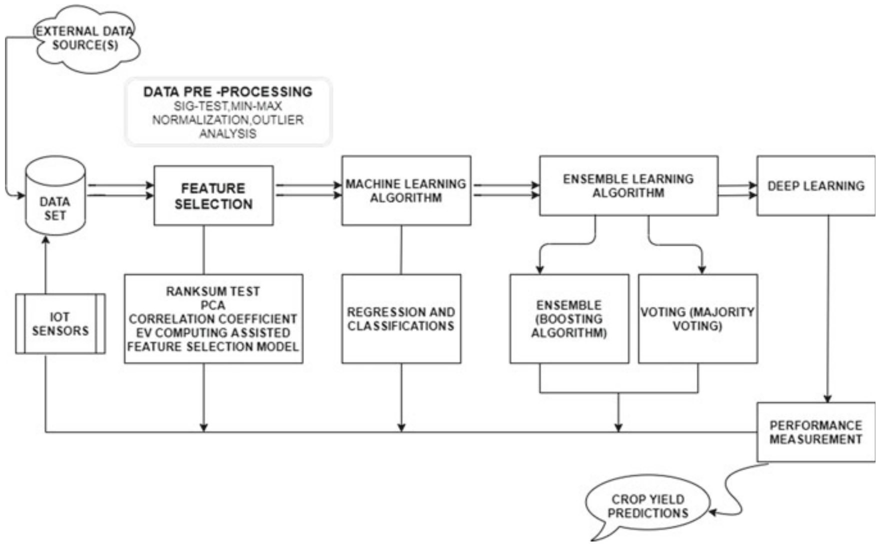


Fig. 2 Proposed model for smart farming prediction and intelligent agriculture analytics model

the heterogeneous dataset. Apart from these, also we need to perform the algorithms (Rank Sum Test, Principle Component Analysis Pearson Correlation Analysis, Evolutionary Computing-assisted feature selection) for feature selection.

Phase#3: Following the data preprocessing processes, we must use a machine learning algorithm to do classification and regression. In this research, two different HELE models, HELE-MVE and HELE-BTE have been applied. Different kinds of machine learning approaches such as decision trees, regression techniques, neural networks, advanced neural networks, enhanced pattern mining models SVM, and SVR have been applied as base classifiers. Eventually, with such heterogeneity, the final prediction outcome can be more efficient and reliable. This approach can achieve more computationally efficient, accurate, and constructive decisions towards optimal smart farming and/or intelligent agriculture. The overall approach, which exploits the MPO concept with an enhanced and computationally enriched model, can be vital in achieving optimal smart farming and/or intelligent agriculture decisions. Thus, the overall proposed model can be significant in achieving optimal smart farming and/or intelligent agriculture decisions.

5.1 Machine Learning Algorithms

A decision tree is a supervised machine learning technique that can be used to solve classification and regression problems. Internal nodes represent dataset properties, whereas branches represent decision rules, and each leaf node represents the conclusion. There are two nodes in a decision tree: the decision node and the leaf node. Decision nodes have multiple branches used to make decisions and, however, leaf nodes are the outcome of those decisions and do not contain any further branches. Based on the features of the given crop dataset decisions or the test are performed. The KNN algorithm is a benchmark classifier that is frequently used to classify more complex data. Despite its simplicity, KNN is a more powerful classifier that is currently employed in the majority of applications. KNN was first used in 2006 to assign genes based on their expression profiles, but it is currently utilized in a variety of sectors for categorization and produces excellent results. The output of KNN classification is a class with related instances. The input given by us will be classified based on vote (majority) of its neighborhoods, considering K-value being assigned to the class.

5.2 Ensemble learning algorithm

This ensemble learning basically consolidates different ML algorithms. It is a hybrid algorithm and with voting mechanism the best ensemble model can be considered. This would increase categorization performance. Using more than one classifier, the ensemble model produces a superior model. The ensemble model increases the accuracy of categorization by producing several classifiers. By summing the outcomes of their classification methodologies, the ensemble model leads to a superior decision-making system.

5.3 Voting

Voting may be a measuring tool for choosing the best ensemble learning algorithms among many alternatives.

5.4 Deep Learning

In addition to the machine learning algorithms proposed, we can also apply below mentioned deep learning techniques such as:

- LSTM

- RNN
- DNN
- DBN
- CNN
- DAE.

Further, with proper analysis and implementation of the ensemble learning model (HELE) we can make our recommended system more efficient and robust.

5.5 *Performance Measurement*

Data were obtained from several IoT sensors and other sources such as CNR and ISTAT, and then preprocessed using the sig test, min–max normalizations, and outlier analysis. The rank sum test, correlation coefficient, and EV computing-assisted features selection model were used for feature selection. The acquired data are then used to run machine learning techniques such as regression and classification and ensemble learning-applied boosting method to identify majority voting and performance assessment. Parameters to be consider to evaluate a machine learning, model's performance are confusion matrix, accuracy, precision, recall, specificity, $F1$-score, precision-recall or PR curve, receiver operating characteristics (ROC) curve. More datasets can be supplied into the system to improve accuracy, and analysis can be performed to guarantee that performance is improved.

6 Conclusion and Future Scope

In our proposed system, we provide an effective recommendation system for better farming by analyzing the previous historical data collected from various sources and live data collected from the different types of sensors. Other ML algorithms and heterogeneous dataset and deep learning techniques may be integrated for better predictions, performance, and user interface, instead of using only historical data collected from very few sensors. After collecting real agricultural data from widely used sensors for better accuracy and output, and efficient recommendation system may be developed to help farmers increase their productivity with very low losses.

References

1. Khaki S, Wang L, Archontoulis SV (2020) A CNN-RNN framework for crop yield prediction. Front Plant Sci 10:1750
2. Agarwal S, Tarar S (2021) A hybrid approach for crop yield prediction using machine learning and deep learning algorithms. J Phys Conf Ser 1714(1)

3. Wang X et al (2020) Winter wheat yield prediction at county level and uncertainty analysis in main wheat-producing regions of China with deep learning approaches. Remote Sens 12(11):1744
4. Alibabaei K, Gaspar PD, Lima TM (2021) Crop yield estimation using deep learning based on climate big data and irrigation scheduling. Energies 14:3004
5. Gong L et al (2021) Deep learning-based prediction on greenhouse crop yield combined TCN and RNN. Sensors 21(13):4537
6. Khaki S, Wang L (2019) Crop yield prediction using deep neural networks. Front Plant Sci 10:621
7. Bi L, Hu G (2021) A genetic algorithm-assisted deep learning approach for crop yield prediction. Soft Comput:1–12
8. Shahhosseini M et al (2021) Coupling machine learning and crop modeling improves crop yield prediction in the US corn belt. Sci Rep 11(1):1–15
9. Chu Z, Jiong Y (2020) An end-to-end model for rice yield prediction using deep learning fusion. Comput Electron Agric 174:105471
10. Gopal PSM, Bhargavi R (2019) A novel approach for efficient crop yield prediction. Comput Electron Agric 165:104968

A Road Map for Classification of Heart Disease Using Machine Learning Classifier

Sibo Prasad Patro, Neelamadhab Padhy, and Rahul Deo Sah

Abstract Heart disease becomes one of the most influential diseases that cause a large number of deaths every year around the world. A report by WHO shows in the year 2016 nearly 17 million people gets died due to heart disease every year. The death rate is increasing rapidly day-by-day and it is estimated by WHO that this death ratio will reach the peak of 75 million by 2030. Despite the availability of modern technology and health care system, prediction and diagnosis of heart disease are still beyond the limitations. Currently, the clinical industries and diagnosis centers have a huge of amount data for the diagnosis of heart disease patients. Machine learning algorithms are more useful to find the hidden patterns, discover knowledge from the dataset, and predict correct outcomes. This research proposed an efficient machine learning-based classifier methodology that outperforms the existing similar methodologies. To evaluate the proposed machine learning classifier, we have taken data from the UCI repository. In this study, we have used ZeroR, bagging, M5, and decision table classifier. The M5 classifier produced a good result compared to other classifiers with 0.2726 mean absolute errors.

Keywords Heart disease · Classification · Prediction · Feature selection · ZeroR · Bagging · M5 · Decision table

S. P. Patro (✉)
School of Engineering and Technology, Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India
e-mail: sibofromgiet@giet.edu

N. Padhy
School of Engineering and Technology, Department of Computer Science and Engineering, GIET University, Gunupur, Odisha, India
e-mail: dr.neelamadhab@giet.edu

R. D. Sah
Dr. Shyama Prasad Mukherjee University, Ranchi, India

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_59

1 Introduction

Heart disease is rapidly becoming one of the leading causes of death worldwide. Nearly 17 million people die every year due to cardiovascular disease [1]. Since the last three quarters, it has been noticed that the deaths from heart disease are mostly found in the low and middle-income countries, now it is alarming in developed counties. According to the Centers for Disease Control and Prevention, heart disease is responsible for about a quarter of all deaths in the United States. Cardiovascular disease is one of the main causes of death not only in the United States, but also in many other nations throughout the world, affecting men and women of all ages, races, and socioeconomic groups. Though the rapid development of technology and the tremendous progress of medical science over the world, preventing different types of heart disease is yet to possible. The death rate of heart disease is increased by 35% in Bangladesh between 1986 and 2006 [2]. The majority of people develop heart disease in their early years of life [3]. In India, 50% of people are suffering from heart disease before 50 years, whereas 25% of people are suffering the same before 40 years. Though heart disease is creating a catastrophic moment for both patients and health authorities, still vital challenge is to predict and detect its presence in the human body despite having different techniques [4].

Today a large number of medical data or records are available with us, but these data are from different foundations. The datasets are required in-depth analysis by the physicians. The raw data might consist of some noise, inconsistency, and incomplete. Before using these data in machine learning for disease prediction in-depth data, preprocessing is required to fill the null values in the dataset. For identifying heart disease in the human body, a correct methodical tool is required, and mining the data for timely analysis of heart infection is a serious requirement. The physicians maintain a huge set of patient records and store them. It helps the physician a great hope for mining valued knowledge from these datasets. It is a challenging job for predicting heart disease and it should be taken with most care. The heart disease prediction algorithms should ensure the accuracy of the results and the generalization of the algorithms should be high. That is why, to decrease the death rate and protect every family from economic vulnerability, prediction of heart disease is an important factor that will ultimately help policymakers to take appropriate steps anent heart disease.

Machine learning is one of the useful instruments that help in a large number of health technology datasets, it helps increasing access and analysis of healthcare facilities. A classifier in machine learning is an algorithm that automatically categorizes data into one or more sets of classes. Data mining techniques are used to extract existing relationships and hidden patterns from a large data source using different statistical analyses, database technology, and machine learning [5]. Machine learning classifier and data mining techniques are mostly used for automatic diagnosis with a satisfactory outperformance in the field of the healthcare industry to save patient life. Machine learning algorithms help the stakeholders to take the right decision also help the physician to give the right decision to treat the patients. As a result, in our research, we used machine learning classifier algorithms to predict heart disease.

2 Significance of the Study

The lifestyle of modern people has a significant impact on their lives. People all over the world are suffering from heart-related ailments as a result of their modern lifestyle, stress, habits, and some genetic factors, regardless of their age. The main objective of this paper is to predict the presence of heart disease in the early stages of the human body with high accuracy. This could help the physicians to take proactive measures to control the lives of many people. The classification model is one of the parts of machine learning techniques that help to find the presence of heart disease. In this paper, ZeroR, bagging, M5, and decision table novel machine learning classifiers are proposed for heart disease prediction. The proposed technique is also compared with various existing machine learning classifiers. The primary goal of this study is to improve forecast accuracy. For this research, a cardiac dataset is collected from the UCI data repository. The dataset contains 303 records with 14 various features. The outperformance of our proposed model shows better accuracy than the previously proposed models.

The main contribution of this paper designs a heart disease prediction model that predicts the presence of heart disease in humans as accurately as possible. It is also called the coronary heart disease prediction model. The following step shows the novelty of our model.

I. The heart disease dataset contains some null or missing values and that may lead to reduce the performance of our model, hence using the feature selection technique, the null values or missing values are skewed to reduce the number of features, minimizing the dimensionality of the dataset and increasing the interpretability.
II. Following the data preprocessing stage, a variety of machine learning techniques are used to improve the testing rate and shorten the execution time.
III. Finally, novel machine learning classifiers are evaluated on the training dataset to predict the presence of heart disease in humans as much as accurately.

The rest of the paper is organized as follows: Sect. 3 describes the literature review, Sect. 4 shows the background of the work, and Sect. 5 illustrates the proposed model followed with a demonstration of the results and conclusion with feature scope in Sects. 6 and 7, respectively.

3 Literature Review

One of the world's most significant and hazardous diseases is heart disease. It becomes an important issue and common problem in the world. Thousands of people died of heart disease every year. As a result, many scientists are attempting to forecast the onset of cardiovascular illness. Cardiovascular illness has emerged as a major concern in the field of clinical data analysis. Mohan et al. [6] proposed a hybrid random forest with a linear model combining characteristics and many classification approaches. The model outperformed with an accuracy of 88.7%. Yadav et al. [7] highlighted that heart disease mostly affects parts of the body such as the pulmonary artery, enzaina, and atalata. Contraction or blockage of blood arteries in the heart causes heart disease. The disease affects the human body tissues called pericarditis and these tissues are closest to the heart. Machine learning algorithms offer a method for identifying dataset attributes as well as the relationships between them. Heart disease-related information is collected from the UCI repository for this research that contains 1025 instances with 14 attributes. Pearson correlation, recursive features elimination, and lasso regularization are utilized to choose features in this study. Along with these few classification, algorithms named random tree, M5P, and random forest ensemble methods were used after the feature selection of heart patient's dataset. The results of the proposed model produced 93.4%, 95.2%, 96.6% for M5PT, RT, and REPT, respectively. Chowdhury et al. [8] suggested an approach for heart disease prediction that aims to uncover relevant features using multiple machine learning algorithms to increase the system's accuracy. They gathered data for the planned project by physically visiting various hospitals and healthcare enterprises in Bangladesh's Sylhet district. The dataset consists of 564 instances with 18 attributes. Decision tree, Naive Bayes, support vector machine, and KNN classification techniques were used to train the dataset. With a 91% accuracy level, the SVM delivered the best results. Kausar et al. [9] described, nowadays, a huge amount of biological data related to heart disease are collected from healthcare industries and these collected data are in irregular form. As a result, locating valuable information in a reasonable amount of time and at a reasonable cost is extremely challenging. The irregular data will fail to accurate prediction of heart disease. To overcome such a problem, we need dimensionality reduction feature methods to process the data. They employed LLE as a feature selection approach on the Cleveland and Statlog datasets and then applied decision tree, support vector machine, decision tree, and NN classification methods following the feature selection process. With an AUC-ROC value of 80% for training datasets, the random forest was found to be the best classifier for the prediction of cardiovascular disease.

Salhi et al. [10] researched heart disease from the data analytics point of view. In this research, data analytics techniques including neural network, SVM, and KNN were used to identify the heart disease on a different size of dataset. The structured data were gathered from Algerians who conducted analyses at the Mohand Amokrane EHS Hospital ex CNMS in Algiers, Algeria. The dataset contains a size of 1200 rows and 20 columns. The model was evaluated on the dataset with a preprocessing phase by selecting the most relevant features with the help of a correlation matrix. The neural network produced the outperformance with 93% of accuracy on the proposed model. Amudhini et al. [11] proposed artificial neural network backpropagation method, and the researcher suggested a system for predicting cardiac illness. The training of the dataset was done by a backpropagation algorithm. The performance of the model improved by reducing the number of neurons of the input layer so that the size of the hidden layer is reduced. The model was produced with 95% accuracy. Vivekanandan et al. [12] discussed, data processing and analysis is one of the challenging job on enormous data. In a given dataset, selecting critical features from an enormous set of features is a challenging task, hence the feature selection process is used for data preprocessing in classification problems. For the proposed research, the medical dataset is collected from the UCI repository. The dataset consists of 300 records and 13 attributes. Out of 13 attributes, only 9 critical attributes were taken and the attributes are handled with a modified DE algorithm. To perform feature selection for heart disease prediction, a modified differential evolution (DE) algorithm was presented. The fuzzy AHP and feed-forward neural network with the specified features are used to predict heart disease. The outperformance of the proposed model was produced with an accuracy of 83%.

Pandian et al. [13] proposed an ensemble technique by combining Naïve Bayes and random forest algorithms to improve the performance of classification algorithms. A Cleveland hospital in the United States was used to test the proposed ensemble technique. There are 337 records in the collection, each containing 13 characteristics. The dataset was evaluated with Gaussian Naïve Bayes algorithm and random forest classifier, and the accuracy is produced with 87% for each. The accuracy climbed to 89% when they merged both methods and employed a voting classifier with correct weights. Sharmila et al. [14] suggested a nonlinear classification technique for predicting heart disease accurately. The suggested system uses big data tools including Hadoop Distributed File System (HDFS) and MapReduce, as well as support vector machine, to predict heart disease using a small number of variables. The suggested work used HDFS to investigate various data mining strategies for storing massive data in multiple nodes and evaluating using SVM in multiple nodes at the same time. The performance of the proposed model using SVM produced an efficient accuracy of 82.35%, and when the SVM was used in a parallel fashion the accuracy increased to 85%. Anbuselvan et al. [15] analyzed various supervised learning models including K-nearest neighbors, Naïve Bayes, support vector machine, decision tree, random forest, and XGBoost. A Cleveland heart dataset was obtained from the UCI machine learning repository for examination of the methods. There are 14 attributes and 303 instances in the dataset. Eighty percent of the data was categorized as training and twenty percent as testing. Individual classifiers were used to learn the training dataset. With an accuracy of 86.89%, the random forest exceeded.

According to the results of the literature review, there are still certain difficulties that need to be addressed. Choosing appropriate and effective regulations from among

Table 1 Comparison between proposed model and the state of art

S. No.	Author	Year	Model	Machine learning algorithms used	Accuracy (%)
1	Mohan et al.	2019	A linear model	Hybrid random forest	88.70
2	Yadav, D. C., et al.	2020	Heart disease classification model	Pearson correlation, recursive features elimination, and lasso regularization	96.60
3	Chowdhury, M. N. R.	2021	Heart disease classification model	Decision tree, Naive Bayes, support vector machine, and KNN	91.00
4	Kausar, N., et al.	2020	Heart disease prediction using dimensionality reduction feature technique	Decision tree, support vector machine, decision tree, and NN	80.00
5	Salhi, D. E., et al.	2020	Heart disease prediction model	Neural network, SVM, and KNN	93.00
6	Amudhini, V. P., et al.	2018	Artificial neural network backpropagation model	Artificial neural network	95.00
7	Vivekanandan, T., et al.	2017	E-healthcare model	DE, neural network	83.00
8	Pandian, A., et al.	2018	An ensemble technique	Naïve Bayes and random forest	87.00
9	Sharmila, R., et al.	2018	Hadoop distributed file system (HDFS) and map reduce	Support vector machine	85.00
10	Anbuselvan, P. et al.	2020	Supervised learning models	K-nearest neighbors, Naïve Bayes, support vector machine, DT, RF, and XGBoost	86.89

them is a difficult task. It has also been identified that to predict heart disease selecting the best feature is very important (Table 1)

4 Methodology

The prediction of cardiac disease is one of the areas where machine learning can be employed. To predict heart diseases correctly and more accurately, there are several

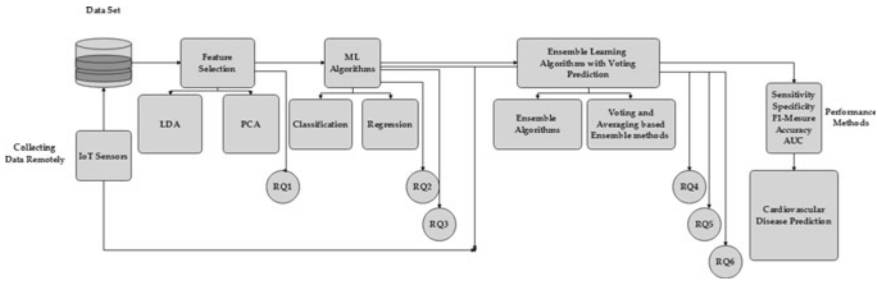


Fig. 1 Detailed research questions with heart disease prediction model

questions raised from data collection to heart disease prediction. Figure 1 shows the detailed research questions in each phase of the heart disease prediction model.

5 Research Questions and Their Solutions

RQ1 Can it be possible to monitor an accurate heart rate for an aged person?

Solution: Machine learning is a new approach that aids in disease prediction and diagnosis. With the rapid advancement of technology and data, the healthcare domain is one of the most significant study fields in the contemporary era. The enormous amount of patient data is tough to manage. Big data analytics makes it easier to manage this information. The data can be collected using various IoT sensors. The researcher proposed an IoT-blockchain-based health monitoring system model [16] through which the patient's automatic heart rate can be calculated and the heart condition can be predicted. The following step shows how the model helps to monitor an accurate heart rate for an aged person.

I. The patient health data are collected from a wearable device and the data are sent to blockchain for feature operations.
II. Classify heart rate based on health and age using the Bayesian network.
III. Predict the heart condition using a multilayer perception neural network.
IV. Send the feedback to the user for precaution.
V. An automated health monitoring system includes quantifying a patient's heartbeat rate values, different ways for collecting the heartbeat rate of the patient, future prediction of heartbeat rate counting, and a proposed machine learning method for heart disease classification.

Machine learning (ML) is used effectively in assisting the making of a decision and prediction. The patient health data are collected from a wearable device and how the data storing in blockchain for feature operations. Machine learning nowadays is used in various areas of the IoT. Multiple studies give a glimpse into a prediction of heart-related disease with the help of machine learning techniques. For the RQ1, we have proposed an automated health monitoring system including quantifying patient's heartbeat rate values, different ways for collecting the heartbeat rate of the patient, future prediction of heartbeat rates counting, and a proposed machine learning method for heart disease classification using Bayesian network.

RQ2 Can it be possible to detect heart issues at an earlier stage and intimating the patient to improve their health using machine learning algorithms?

Solution: The healthcare business and researchers are confronted with difficult concerns as a result of the aging population and challenges in accessing health care. An unhealthy diet, physical inactivity, cigarette use, and alcohol consumption are the leading causes of cardiovascular disease and stroke. Behavioral risk factors may affect those who have high blood pressure, high blood glucose, high blood lipids, and are overweight or obese. Even in certain circumstances, the heart attack strikes without warning. It is a challenging burden for doctors to keep track of their patients' states in real time, and it is one of the leading causes of mortality. To address these issues, the authors suggested a framework based on several machine learning classifiers for predicting heart disease utilizing primary risk factors [17]. K-nearest neighbor, Naive Bayes, support vector machine, lasso, and ridge regression techniques were proposed in this study. The suggested model was tested using data from the UCI machine learning repository on Cleveland heart disease. Linear discriminant analysis and principal component analysis were utilized for data classification and feature selection. The presence of cardiovascular disease can be detected with 92% accuracy using the support vector machine and 85% accuracy using the $F1$ algorithm. Precision, accuracy, and sensitivity are used to assess the planned research work's outperformance.

RQ3 Can we calculate the accuracy of heart disease at an early stage by using a novel optimization algorithm to minimize error, cost, using various machine learning classification algorithms?

Solution: Analysis of data in the healthcare domain becomes a challenging job. Data analysis is a critical component in confirming the diagnosis, fine-tuning the research methodologies, and choosing relevant equipment based on the diseases' requirements. Artificial intelligence and machine learning are used to examine current data to make the best predictions possible. From test and training data, a prediction model can be created using the classification method. Through a mix of mathematical tools and computer-aided procedures, the testing and training data are screened by a classification algorithm, resulting in a new model capable of detailed data and containing the same classes of data. We utilize the optimization technique to examine current data and forecast ideal outcomes. Optimization algorithms offer the benefit of being able to cope with complex nonlinear issues while remaining flexible and adaptable. Optimization is the process of identifying a set of inputs to an objective function that results in a maximum or minimal function evaluation. Many machine learning algorithms, from fitting logistic regression models to training artificial neural networks, are based on this difficult topic. To address this issue, we suggested a framework for predicting heart disease based on main risk variables and a variety of classifier algorithms, including Naive Bayes (NB), Bayesian optimizedsupport vector machine (BO-SVM), K-nearest neighbors (KNN), and salp swarm optimizedneural networks (SSA-NN) [18]. A heart disease dataset was obtained from the UCI machine learning repository to test the system. Different classifier approaches are proposed in this work, including the use of a combination of ensemble-based machine learning algorithms to identify redundant characteristics to improve the accuracy and quality of heart disease classification. The suggested study's goal is to determine the greatest accuracy for predicting heart disease utilizing main risk variables and several classifier algorithms such as Bayesian optimized support vector machine (BO-SVM), K-nearest neighbors, and others (KNN). The BO-SVM model outperformed SVM (accuracy = 93.3%, precision = 100%, sensitivity = 80%), followed by SSA-NN (accuracy = 86.7%, precision = 100%, sensitivity = 60%). The suggested framework's results show that a novel optimized algorithm can provide an effective healthcare monitoring system for early cardiovascular disease prediction.

RQ4 Can it be possible to improve the prediction of cardiovascular disease using ensemble learning techniques by applying majority voting prediction?

Solution: The heart is an essential organ in the human body on the off chance this organ gets influenced, and then it equally influences the other fundamental pieces of the body. Making the need for an effective prediction system a source of high demand in treating affected patients. Ensemble methods are a type of machine learning that combines numerous basic models to create a single most productive model [19]. When the best performance on a predictive modeling assignment is the most essential outcome, ensemble learning approaches are popular and the go-to strategy. To analyze and improve the prediction of cardiovascular disease, an automated medical diagnosis system can be developed. An ensemble can make better forecasts and accomplishes better results than a single contributing model. Using an ensemble reduces the spread or dispersion of predictions and model performance. Several powerful techniques can be used, namely Max voting, averaging, weighted averaging, bagging, and boosting. Voting is one of the simplest ensemble algorithms and is very effective for heart disease diagnosis and prediction. This technique is used for classification or regression problems. Initially, it helps to create two or more sub-models. Each sub-model makes a prediction which is combined using majority voting rules. It is also known as a meta-classifier because it uses majority vote to combine different machine learning classifiers for categorization. Majority voting aids in the prediction of final class labels by identifying the class label that has been predicted the most frequently by classification models.

RQ5 How to identify the weakness of one or more classification algorithms and use ensemble methods to boost their performance?

Solution: Many researchers proposed various intelligent e-healthcare decision support systems using various machine learning algorithms to improve the performance of cardiovascular disease. These proposed systems help to identify the presence of cardiovascular disease in the human body at an early stage and in time. But in those systems, a single classifier algorithm was used and the accuracy was not satisfactory. To overcome the issue, we can implement a new technique called hybrid ensemble machine learning algorithms by building a model for the prediction and detection of CVD at an early stage in the human body [20]. The suggested boosting ensemble learning algorithms will aid in the transformation of a group of weak learners into strong learners, reducing training errors. In this kind of approach, sample of data is selected randomly then fitted with the model and then trained sequentially. In this process, each model tries to compensate for the weaknesses of its predecessor.

RQ6 Can it be possible to use novel machine learning classifiers for heart disease?

Solution: One of the most pressing difficulties and challenges in clinical data analysis is the prediction of cardiovascular disease. Machine learning is critical for generating decisions and predictions based on the massive amounts of data generated by hospitals and the healthcare industry. We suggest a narrative method for identifying key aspects using machine learning algorithms, which improves the accuracy of cardiovascular disease prediction. A prediction model is proposed that incorporates a variety of characteristics and classification algorithms. In this study, we used the ZeroR, bagging, M5-rules, and decision table algorithms to improve the quality of heart disease classification by filtering out duplicate features. The findings reveal that the suggested method outperforms the proposed method in terms of efficiency and robustness in processing various types of data for heart disease classification when applied to the heart disease dataset. The correlation coefficient, mean absolute error, root mean squared error, relative absolute error, root relative squared error, and other performance indicators are used to evaluate the output of various machine learning classifiers.

5.1 Proposed Model for Heart Disease Prediction

This section contains our feature classification and selection system. The suggested work explores the above-mentioned four algorithms and performs performance analysis to assist forecast the presence of heart disease more accurately. The main goal of this study is to effectively and correctly predict if the patient suffers from heart disease. The Heart disease dataset is provided to evaluate by the proposed model. The data are provided into the model and the model predicts the probability of having heart disease. Figure 2 shows the detailed process involved.

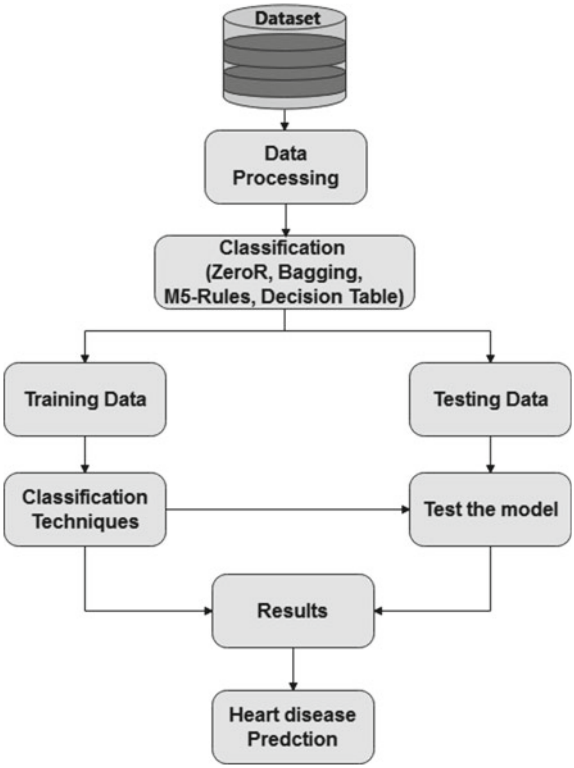


Fig. 2 Proposed model for heart disease prediction

Our proposed system contains four different feature selection-based algorithms including ZeroR, bagging, M5-rules, and decision table algorithms. The training dataset is produced using a binary class classification, features are taken from the training set, and the model is tested using the best features.

5.2 Dataset and Attributes

The heart disease dataset for our proposed heart disease prediction algorithm was obtained from the UCI data repository. The UCI machine learning repository has a large and diverse collection of datasets, including datasets from a variety of areas. The dataset contains 303 records with 14 medical features. The reason of choose this dataset is, it has required no features and is mostly required for heart disease prediction. It has fewer missing data and is commonly used in scientific research [21].

5.3 Data Preprocessing

The dataset contains few NaN values, as the NaN values cannot process and will lead to reduced performance and accuracy, hence these values need to convert into numerical values. In this process, mean of the attribute is calculated and the NaN values are replaced by the mean.

5.4 Splitting

In our paper, the dataset is split into training and testing datasets. About 80% of the data is taken for training and the other 20% is taken for testing.

5.5 Classification Task

Heart disease prediction can be viewed as a clustering or classification problem in the context of machine learning. The issues can be decreased by classification because the model is developed for a large set of presence and absence of data. The challenge can be reduced to one classification study, which has a limited set of classes, including the heart disease sample, for the known families. Clustering methods make it easier to identify the correct class and improve performance accuracy. In this section, the theoretical context for all of the methodologies employed in our research is offered. For the comparative analysis, we have used four different algorithms. The popularity of these algorithms was the primary factor in their selection [22].

5.6 Feature Selection

In the collected dataset, 14 attributes are available. Because these expression data contain a variety of useless and repetitive features, the process of classifying heart disease becomes increasingly difficult. If all of the data is used to classify heart disease, the performance will be diminished and the accuracy will be compromised. Hence, the feature selection approach can be applied for reducing the dataset size, eliminating unwanted data, and improving the performance and accuracy.

The detail of the four different algorithms used, i.e., ZeroR, bagging, M5-rules, decision table are explained in detail.

5.6.1 ZeroR Classification Algorithm

It is based on the aim and disregards all predictors. It merely forecasts which category will be in the majority. This method is particularly beneficial for determining baseline performance as a comparison point for the other classification methods. The output is simply the most frequently identified classification for a given set of data. If we say, 65% of data items have the classification, then ZeroR classification algorithm identifies all the data items are having in it and the performance would be right as 65% of the time.

5.6.2 Bagging Classification Algorithm

It is a meta-estimator for ensembles. It uses a random subset of the original dataset to fit the basic classifiers, and then aggregates their predictions to create a final prediction. This type of meta-estimator is typically used to lower the variance of a black-box estimator by removing randomness from the creation process and then removing ensemble from the equation. On a randomly generated dataset, all the base classifiers are trained parallel. Each base classifier's training set is distinct from the others. The original data is iterated in the training set those results, and the rest is discarded. Moreover, in our proposed work, we applied tenfold cross-validation.

5.6.3 M5-Rules Classification Algorithms

It generates a decision list for regression problems using a separate and conquer. In this approach, through each iteration, it constructs a new model tree and generates the best leaf into a rule. For regression tasks, the M5 model tree, also known as the decision tree learner, is commonly used to predict numerical response variables. It helps to predict continuous numerical attributes. In our proposed model for heart disease prediction, we applied tenfold cross-validation on the dataset and using smoothed liner models applied seven rules.

5.6.4 Decision Table Classification Algorithms

DT algorithms are a short visual representation for identifying which task to perform based on the given condition. In these algorithms, output is a set of actions. The information which is used in the decision table can be defined as a decision tree. In this approach, the given decision interacts with a variable whose possible values are listed among a set of alternative conditions. The specified operation specifies whether the action is to be performed or not.

Table 2 Classification report

Classifier	Correlation coefficient	Mean absolute error	Root mean squared error	Relative absolute error (%)	Root relative squared error (%)
ZeroR	0	0.496	0.498	100	100
Bagging	0.678	0.283	0.366	56.951	73.430
M5	0.665	0.272	0.373	54.844	74.913
Decision table	0.574	0.274	0.419	55.269	84.071

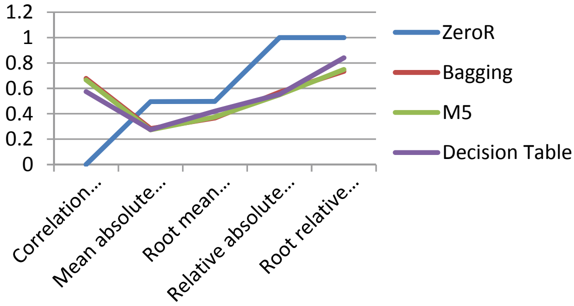


Fig. 3 Comparison with different classified algorithms

6 Experimental Analysis

After a set of analyses, the experimental results were obtained. In this work, we have used test dataset of about 20% and the training dataset with 80%, including a total of 100% data is utilized in this experiment. The dataset is classified according to the patient's heart characteristics in different structured data. Table 2 shows the different accuracy from the test dataset.

From the above table, it shows that the M5 classifier mean absolute error produced with 0.2726 and which is better compared to other classifiers that are used in our model. The graphical representation graph is shown in Fig. 3.

7 Conclusion

Heart disease becomes one of the crucial health problems all over the globe. A systematic and scientific prediction approach can reduce the loss of heart disease. Using a variety of machine learning classifiers, we developed a method for predicting cardiac disease. In our paper, we have used ZeroR, bagging, M5-rules, decision table classifier. The result of our study shows that the mean absolute error for ZeroR, bagging, M5-rules, and decision table are 0.496, 0.2831, 0.2726, and 0.2748, respectively. Among the classifiers, we have identified that M5-rule produced the mean absolute error with only 0.2726. The relative absolute error of 100% and root relative squared error of 100% were also achieved using the ZeroR classifier. As per the results we approach our proposed model can be the best model for heart disease classification. In our future work, we will be using the correlation between the different attributes in the dataset.

References

1. World Health Organization/health-topics/cardiovascular-diseases
2. Ahsan Karar Z, Alam N, Kim Streatfield P (2009) Epidemiological transition in rural Bangladesh, 1986–2006. Glob Health Action 2(1):1904
3. Mittal RA (2017) Increasing heart attacks in young Indians. The Times of India
4. Chandra P, Deekshatulu BL (2012) Prediction of a risk score for heart disease using associative classification and hybrid feature subset selection. In: 2012 12th international conference on intelligent systems design and applications (ISDA). IEEE, pp 628–634
5. Dhar S, Roy K, Dey T, Datta P, Biswas A (2018) A hybrid machine learning approach for prediction of heart diseases. In: 2018 4th international conference on computing communication and automation (ICCCA). IEEE, pp 1–6
6. Mohan S, Thirumalai C, Srivastava G (2019) Effective heart disease prediction using hybrid machine learning techniques. IEEE Access 7:81542–81554
7. Yadav DC, Pal SAURABH (2020) Prediction of heart disease using feature selection and random forest ensemble method. Int J Pharmaceutical Res 12(4):56–66
8. Chowdhury MNR, Ahmed E, Siddik MAD, Zaman AU (2021) Heart disease prognosis using machine learning classification techniques. In: 2021 6th international conference for convergence in technology (I2CT). IEEE, pp 1–6
9. Kausar N, Ghous H (2020) A novel framework for classification of cardiovascular disease using local linear embedding as feature selection method. LC Int J STEM 1(3):17–31. ISSN: 2708–7123
10. Salhi DE, Tari A, Kechadi MT (2020) Using machine learning for heart disease prediction. In: CSA, pp 70–81
11. Amudhini VP, Santhini T, Kailash P, Nivetha D, Poonguzhali R (2018) Survey on machine learning algorithms for prediction of cardiovascular disease
12. Vivekanandan T, Iyengar NCSN (2017) Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease. Comput Biol Med 90:125–136
13. Pandian A, Manikandan K, Vamsi NK, Aditya BMSSS. Prediction of cardiovascular diseases using hybrid machine learning
14. Sharmila R, Chellammal S (2018) A conceptual method to enhance the prediction of heart diseases using the data techniques. Int J Comput Sci Eng 6(4):21–25
15. Anbuselvan P (2020) Heart disease prediction using machine learning techniques. Int J Eng Res Technol (IJERT):2278-0181
16. https://www.routledge.com/The-Role-of-IoT-and-Blockchain-Techniques-and-Applications/Kuanar-Mishra-Peng-Dasig-Jr/p/book/9781771889148
17. Patro SP, Padhy N, Chiranjevi D (2021) Ambient assisted living predictive model for cardiovascular disease prediction using supervised learning. Evol Intel 14(2):941–969
18. Nayak GS, Patro SP, Padhy N (2021) Heart disease prediction by using novel optimization algorithm: a supervised learning prospective
19. Mienye ID, Sun Y, Wang Z (2020) An improved ensemble learning approach for the prediction of heart disease risk. Inform Med Unlocked 20:100402
20. Sammout R, Salah KB, Ghedira K, Abdelhedi R, Kharrat N (2020) A proposal of clinical decision support system using ensemble learning for coronary artery disease diagnosis. In:

International conference on wireless mobile communication and healthcare. Springer, Cham, pp 300–314

21. U. H. Dataset. UCI machine learning repository [online]. https://archive.ics.uci.edu/ml/machine-learning-databases/heartdisease/Hear
22. Orphanou K, Stassopoulou A, Keravnou E (2015) DBN-extended: a dynamic Bayesian network model extended with temporal abstractions for coronary heart disease prognosis. IEEE J Biomed Health Inform 20(3):944–952

An Effective Optimization Method for Encroacher Detection System Using Deep Learning Technology

S. V. S. V. Prasad Sanaboina and K. Rajiv

Abstract Encroacher detection machine ought to obtain proper precise and adaptableness to retribution attacks from encroacher. Encroacher detection machine distinguishes among the valid and criminal users and ought to be used with the primary line of security to thwart encroachers and aberrations from inner in addition to outdoor attackers. The encroacher detection machine is a critical asset to PC security, due to the fact that attacker attempts to hide his/her identity, and release attacks through intermediate hosts broadly called stepping stones encroacher. Secondly, converting creation of generation and approach makes it extra hard to come across attacks. These cannot obtain by existing methods, those various community attacks are modeled via way of means of using different types of traffic, that is, expensive and complex. The initiate encroacher detection machine can consequently employ optimization algorithms strategies in gadget studying to come across the unknown destiny assaults. In the exploration, a brand-new version is initiate for detecting the future encroacher attacks. Experimental result confirmed that the improved encroacher detection model mixed with deep notion network efficaciously improves the popularity price of encroacher attacks and reduces the complexity of the network. The adequacy of the start approach is set up contrary to current systems through method-of-method for the utilization of following measurements, for example, Accuracy, Sensitivity, Specificity, Precision, *F*-score, and Mathew's connection coefficient.

Keywords Encroacher detection system · Attracts · Optimization algorithms · Metrics · Deep learning

S. V. S. V. Prasad Sanaboina
GIET University, Gunupur, Odisha, India
e-mail: prasad.sanaboina@mriet.ac.in

K. Rajiv (✉)
GRIET-JNTUH, Hyderabad, India
e-mail: rajiv.research31@gmail.com

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6_60

1 Introduction

Traditionally, encroacher is detected through authentication, encryption and decryption techniques, firewalls, etc. These are called first line protection in computer protection in which it permits assessment of computer applications installed at the host to detect the known vulnerabilities. After assessment, the susceptible computer application is patched with the present-day patch code [1]. However, attacker can pass them effortlessly and primary line protection mechanism is not bendy and effective sufficient to thwart different forms of attacks or encroachers [2, 3]. So, antivirus softwares that act as the second one line of protection, which has a limitation, and it can most effective hit upon attacks, whose segregations are gift in the database. Antivirus softwares are restricted of their ability to deal with attacks that could affect in some hours till their subsequent updates. However, the strong creation of encroacher detection system gathers records related to activities that violate security policies [4]. Encroacher detection system gathers records from a network system and examines to decide the segments, which violate protection policies of PC and networks. Accuracy, extensibility, and adaptability are the three important characteristics of encroacher detection system (Fig. 1).

Encroacher detection system needs to gain good accuracy and flexibility to counter attacks from intruders. Encroacher detection systems distinguish among the valid and illegitimate users and need to be used with the primary line of protection to thwart encroachers and aberrations from inner in addition to outside attackers.

Common Types of Cyber-Attacks

See Fig. 2.

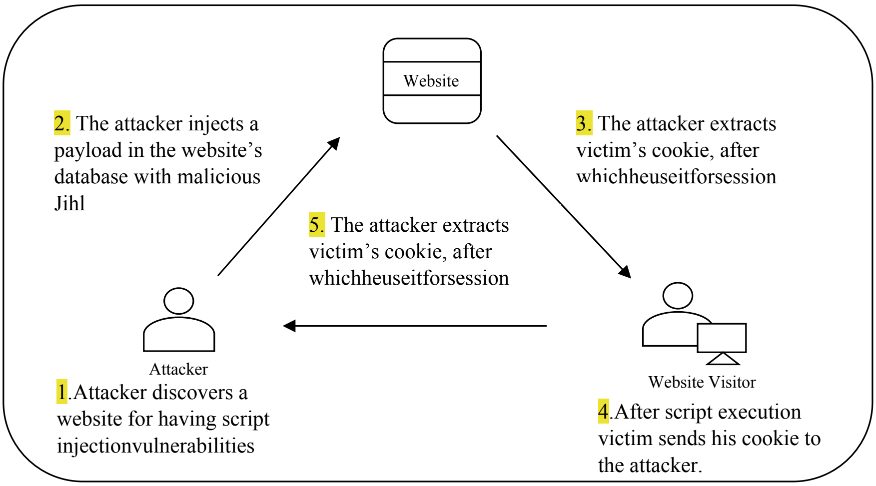


Fig. 1 How to attacker attack the website

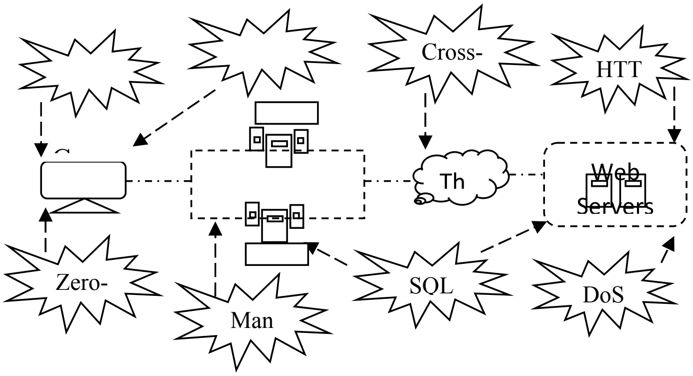


Fig. 2 Common types of cyber-attacks

2 Literature Survey

Manzoor and Kumar [5] developed an intelligent system, which first performs feature ranking based on information gain and correlation.

Then, feature reduction is carried out via way of means of combining ranks obtained from each data benefit and correlation via way of means of utilizing a singular approach to become aware of beneficial and vain features. These decreased capabilities are then fed to a feed ahead neural network for schooling and trying out on KDD99 dataset. In this study, preprocessing of KDD99 dataset is carried out to normalize quantity of times of every elegance earlier than schooling. So, the machine behaves intelligently to classify check data into assault and non-assault classes. The intention of the function decreased machine is to gain identical diploma of overall performance as an ordinary machine. The machine is examined on five specific check datasets and each person and common outcomes of all datasets are reported, which might be significant in comparison with the existing models.

Ali et al. [6] fostered a learning model for fast learning network (FLN) in light of particle swarm optimization (PSO) for encroacher discovery. The created model is applied to the issue of encroacher identification and approved dependent on the renowned dataset KDD99. The created model has been analyzed against a wide scope of meta-heuristic calculations for preparing ELM and FLN classifier. The PSO-FLN has beated other learning approaches in the testing precision of the learning.

Kasongo and Sun [7] encouraged a feed forward deep neural network far-off encroacher revelation system using a wrapper-based feature extraction unit 9WFEU. The extraction procedure for the WFEU utilizes extra trees estimation to create diminished ideal component vector. The ampleness of WFEU-FFDNN is pondered subject to the UNSW-NB15 and the AWID encroacher revelation datasets. Also, the WFEU-FFDNN is differentiated and standard AI computations; discretionary forest, unsuspecting Bayes, decision tree, and k-nearest neighbor. The results suggested that the WFEU-FFDNN has more unmistakable revelation precision than existing systems.

Xu et al. [8] fostered an original encroacher location framework that comprises a repetitive neural organization with gated intermittent units, multi-facet insight, and softmax module. Trial on the KDD99 and NSL-KDD datasets showed that the created framework got better execution in encroacher discovery. Near tests showed that gated intermittent units are more reasonable as a memory unit for encroacher recognition framework than long momentary memory organization and demonstrated that it is a viable disentanglement and improvement of long transient memory organization. Subsequently, bidirectional gated repetitive units can arrive at the best exhibition contrasted with distributed techniques.

Zhang et al. [9] fostered an encroacher recognition model based on worked on hereditary calculation and profound conviction organization. Confronting various kinds of assaults through different cycles of hereditary calculation is diminished by adaptively choosing ideal number of stowed away layers and number of neurons in each layer. In this way, the encroacher discovery model dependent on profound conviction network accomplishes a higher recognition rate. Ultimately, the NSL-KDD dataset is used to reenact and assess the model calculation. Test result showed that the further developed encroacher location model joined with profound conviction network adequately further develops the acknowledgment pace of encroacher assaults and diminishes the intricacy of the organization.

3 Initiate Work

In the exploration, a new model is initiated for detecting the future encroacher attacks, i.e., "chicken swarm optimization algorithm". Actually this concept not available in machine learning; this is the new concept to detecting the future attacks. In the exploration study, we are using ANN primarily-based totally encroacher detection which is promising for reducing the number of false negative or false positive due to the fact that ANN has the functionality of getting to know from real examples. Developed learning model for fast learning networks based on chicken swarm optimization (CSO) has been initiated.

In the exploration study, NSL-KDD dataset is used for experimental investigation. The most important task in this experimentation is the selection of dataset, because system performance can be calculated by using the correctness of a dataset. The effectiveness of the system is greater, when the more accurate data are selected. The collection of datasets is based on four methodologies, namely simulated dataset, tested dataset, sanitized dataset, and standard dataset. However, there are more complications occurred in the first three methodologies. Hence, the designing of simulated dataset is highly complex and challenging. Sanitized dataset is not safe to use, where real traffic method is expensive. In addition, a various network attacks are modeled by using different types of traffic, which is expensive and complex.

NSL-KDD dataset is used to overcome these difficulties and to validate the developed method in encroacher detection. There are three parts presented in NSL-KDD dataset: whole dataset, half dataset, and 1/4 dataset, where the standard dataset is randomized. The total of 65,535 samples are presented in the whole dataset, 32,767 samples are included in the half dataset, and 1/4 dataset contains 18,383 samples. Four different attacks are available in the NSL-KDD dataset, which are described below:

- **Denial of Service Attack (DoS)**

A service cannot be provided by the system, due to the increases of network traffic usage or resources of system that is occurred by the results of this attack. The attack types included in DoS are Apache2, Pod, Back, Worm, Teardrop and Process table, Land, Edstrom, and Neptune and Smurf.

- **User to Root Attack (U2R)**

Once an ordinary account is achieved, U2R gets access to the root account. Types of attacks presented in U2R are Load module, Sql attack, Rootkit, Ps, Buffer overflow, Xterm, and Perl.

- **Remote to Local Attack (R2L)**

A user account is obtained by sending a packet to machine and detecting the weakness in the network. A various types of attacks presented in R2L; Ftp_write, warez client, Snmpguess, Named, Spy, Warez master, Guess Password, Send mail, Xsnoop, HTTP tunnel, Imap, Phf, Multichip, Xlock, and Snmpgetattack.

- **Probe Attack**

The network will be scanned for misuse by gathering information about the network weaknesses. Nmap, Mscan, Satan, Saint, Ipsweep, and Portsweep are the attacks presented in Probe attack.

Data Preprocessing

Due to some symbolic characteristics, the classifier cannot process the raw dataset. Therefore, preprocessing is essential that eliminates non-numerical or symbolic characteristics, because they do not play a major role in detecting encroacher. However, this process generates overhead, which includes more training time, architecture of the classifier is complicated, and memory and computing resources are wasted. Therefore, non-numerical characteristics are eliminated from the raw dataset for better performance of encroacher detection. In the exploration study, a preprocessing technique; min–max normalization technique is used for data preprocessing.

Feature Selection

After information pre-handling, include determination is refined by utilizing changed chicken multitude improvement calculation for picking dynamic component esteems for encroacher grouping. The change chicken multitude enhancement deals with the premise of nonlinear idleness weight. This calculation right off the bat utilizes the variety factor in the chicken molecule position refreshing equation and afterward presents the non-direct diminishing weight dependent on a parabola opening upwards in the chicken area update recipe for working on the chicken area and the chicken area. Consequently, the transformation chicken multitude advancement dependent on nonlinear inactivity weight is simpler to get the worldwide ideal arrangement (ideal element esteem) than unique chicken multitude enhancement calculation and other streamlining calculations.

Classification

After obtaining the optimal feature values, classification is carried by support vector machine or artificial neural network. The support vector machines, a supervised machine learning algorithm, which are used for both classification and regression. However, it is mostly applied in classification problems. In support vector machine, every data item are plotted as a point in n-dimensional space (where "n" is indicated as number of features) with the value of each feature being the value of a particular coordinate. Then, classification is performed by finding hyper-plane which differentiates the classes.

The artificial neural network is a feed forward neural network, which is constructed in which every neural network is a feed forward neural network, which is constructed in which every neuron of a layer is conned to next layer neurons. In this scenario, we constructed three layer feed forward neural network. Input layer has n-neurons that are equal to the number of input features. Output layer has five neurons, which are equal to five output classes, i.e., Normal, DoS, U2R, R2L, and Probe.

Performance Evaluation

The effectiveness of the initiate method is validated against existing techniques by using following metrics such as:

- Accuracy
- Sensitivity
- Specificity
- Precision
- F-score
- Mathew's correlation coefficient.

4 Result and Discussion

To approve the created learning form PSO-FLN, substantial difference has been finished with an alternate number of neurons inside the secret layer of FLN and the first ELM. Advance the FLN boundaries to upgrade the IDS precision in our work had been start various calculations including Genetic calculation (GA), Harmony Search Optimization, and Ameliorated Teaching Learning principally-based absolutely improvement. Also, those same optimization algorithms adoptive primarily-based totally ELM to compare with PSO-FLN as proven. Results confirmed that PSO-FLN almost outperformed different learning models irrespective of converting the number of neurons within side the hidden layer. And response, that is, referred to as ID-6.

The encroacher detection machine is an important asset for PC security, because the attacker tries to hide his/her identity and launches the attack through an intermediate host that is widely known as a stepping stone encroacher.

But existing system having following loop poles to identify attacks perfectly.

- No Accuracy
- No Sensitivity
- No Specificity
- No Precision
- Difficult to identify *F*-score
- Not followed Mathew's correlation coefficient
- No Adaptability.

These are the limitations while we follow the existing concept.

Encroacher detection machines are an important asset for PC security because the attacker tries to hide his/her identity and launches attacks through intermediate hosts that are widely known as stepping stone encroachers. Second, the creation of conversion generation and methods makes it more difficult to encounter attacks. These are not available with existing methods. Those various community attacks are modeled by using different types of traffic, which is expensive and complicated.

The preliminary concept to apply encroacher detection in an attempt to deal with misuses and networking attacks in computer was placed forth via way of means of Dorothy E Denning in 1987. The process is applied via way of means of an encroacher detection system. Presently such systems are broadly to be had with variety, points out the overall ineffectiveness and absence of sufficiency supplied via way of means of the existing commercially to be had systems, this brings to mind the need for ongoing exploration on more dynamic encroacher detection systems. In order to execute the process of encroacher detection, there may be a want to identify ongoing or tried encroachers or attacks at the system or network, this identity data encompass data collection, conduct classification, data reducing, and finally reporting and response, that is, referred to as ID [5].

5 Conclusion and Future Enhancement

The RNN-IDS model not just has a solid demonstrating capacity for encroacher discovery, yet in addition has high exactness in both paired and multiclass grouping. Contrasted and customary grouping strategies, like J48, guileless Bayesian, and irregular backwoods, the exhibition gets a higher exactness rate and location rate with low-bogus positive rate, particularly under the assignment of multiclass arrangement on the NSL-KDD dataset. The model can viably work on both the exactness of encroacher discovery and the capacity to perceive the encroacher type. Obviously, later on investigation, we will in any case focus on decrease the preparation time utilizing GPU speed increase, abstain from detonating and evaporating inclinations, and study the grouping execution of LSTM, bidirectional RNNs calculation in the field of encroacher location.

References

1. Yin C, Zhu Y, Fei J, He X (2017) A profound learning approach for encroacher location utilizing intermittent neural organizations. IEEE Access 54:21954–21961
2. Idhammad M, Afdel K, Belouch M (2018) Appropriated encroacher location framework for cloud conditions dependent on information mining strategies. Proc Comput Sci 127:35–41
3. Beshrati E, Naderan M, Namjoo E (2019) LR-HIDS: logistic regression host-based encroacher detection system for cloud environments. J Amb Intell Human Comput 10(9):3669–3692
4. Dhaliwal SS, Nahid AA, Abbas R (2018) Effective encroacher detechtin system using XGBoost. Information 9(7):149
5. Manzoor I, Kumar N (2017) A feature reduced encroacher detection system using ANN classifier. Expert Syst Appl 88:249–257
6. Ali MH, Al Mohammed BAD, Ismail A, Zolkipli MF (2018) Another encroacher identification framework dependent on quick learning organization and molecule swarm improvement. IEEE Access 6:20255–20261
7. Kasongo SM, Sun Y (2020) A profound learning strategy with covering based element extraction for remote encroacher discovery framework. Comput Secur 92:101752
8. Xu C, Shen J, Du X, Zhang F (2018) An encroacher discovery framework utilizing a profound neural organization with gated repetitive units. IEEE Access 6:48697–48707
9. Zhang Y, Li P, Wang X (2019) Encroacher discovery for IoT dependent on worked on hereditary calculation and profound conviction organization. IEEE Access 7:31711–31722

Author Index

R. Kumar et al. (eds.), *Next Generation of Internet of Things*, Lecture Notes in Networks and Systems 445, https://doi.org/10.1007/978-981-19-1412-6