Lecture Notes in Electrical Engineering 733

V. Bindhu
João Manuel R. S. Tavares
Alexandros-Apostolos A. Boulogeorgos
Chandrasekar Vuppalapati *Editors*

International Conference on Communication, Computing and Electronics Systems

Proceedings of ICCCES 2020

Springer

Lecture Notes in Electrical Engineering

Volume 733

Series Editors

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Naples, Italy
Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico
Bijaya Ketan Panigrahi, Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India
Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, Munich, Germany
Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China
Shanben Chen, Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China
Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore
Rüdiger Dillmann, Humanoids and Intelligent Systems Laboratory, Karlsruhe Institute for Technology, Karlsruhe, Germany
Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China
Gianluigi Ferrari, Università di Parma, Parma, Italy
Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain
Sandra Hirche, Department of Electrical Engineering and Information Science, Technische Universität München, Munich, Germany
Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA
Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China
Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
Alaa Khamis, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt
Torsten Kroeger, Stanford University, Stanford, CA, USA
Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA
Ferran Martín, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain
Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore
Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany
Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA
Sebastian Möller, Quality and Usability Laboratory, TU Berlin, Berlin, Germany
Subhas Mukhopadhyay, School of Engineering & Advanced Technology, Massey University, Palmerston North, Manawatu-Wanganui, New Zealand
Cun-Zheng Ning, Electrical Engineering, Arizona State University, Tempe, AZ, USA
Toyoaki Nishida, Graduate School of Informatics, Kyoto University, Kyoto, Japan
Federica Pascucci, Dipartimento di Ingegneria, Università degli Studi "Roma Tre", Rome, Italy
Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China
Gan Woon Seng, School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore, Singapore
Joachim Speidel, Institute of Telecommunications, Universität Stuttgart, Stuttgart, Germany
Germano Veiga, Campus da FEUP, INESC Porto, Porto, Portugal
Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China
Junjie James Zhang, Charlotte, NC, USA

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

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

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

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

China

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

India, Japan, Rest of Asia

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

Southeast Asia, Australia, New Zealand

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

USA, Canada:

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

All other Countries:

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

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

More information about this series at http://www.springer.com/series/7818

V. Bindhu · João Manuel R. S. Tavares ·
Alexandros-Apostolos A. Boulogeorgos ·
Chandrasekar Vuppalapati
Editors

International Conference on Communication, Computing and Electronics Systems

Proceedings of ICCCES 2020

Springer

Editors
V. Bindhu
Department of ECE
PPG Institute of Technology
Coimbatore, India

Alexandros-Apostolos A. Boulogeorgos
Aristotle University of Thessaloniki
Thessaloniki, Greece

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

Chandrasekar Vuppalapati
San Jose State University
San Jose, CA, USA

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-33-4908-7 ISBN 978-981-33-4909-4 (eBook)
https://doi.org/10.1007/978-981-33-4909-4

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

We are honored to dedicate the proceedings of ICCCES 2020 to all the participants, organizers and editors of ICCCES 2020.

Preface

This conference proceedings volume contains the written versions of most of the contributions presented during the ICCCES 2020 Conference. The conference has provided a platform to share and exchange the recent developments in a wide range of topics including computational intelligence, machine learning, signal and image processing, electronic devices and systems, antenna and wave propagation, wireless communication networks and so on. The conference has been a good opportunity for participants coming from various destinations to present and discuss the state-of-the-art topics in their respective research areas.

ICCCES 2020 Conference tends to collect the latest research results and applications on computing, communication and electronics. It includes a selection of 66 papers from 256 papers submitted to the conference from various universities and industries present across the globe. All the accepted papers were subjected to double-blinded peer-reviewing process by 2–4 expert referees. The papers are selected for its high quality and the relevance to the conference.

ICCCES 2020 would like to express our gratitude and appreciation to all the authors for their valuable research contributions to this book. We would like to extend our thanks to Guest Editors **Dr. V. Bindhu, PPG Institute of Technology, India; Dr. João Manuel R. S. Tavares, Rua Doutor Roberto Frias, Porto, Portugal; Dr. Alexandros-Apostolos A. Boulogeorgos, Aristotle University of Thessaloniki, Greece; Dr. Chandrasekar Vuppalapati, San Jose State University, USA;** and all the referees for expressing their constructive comments on all the research papers. In particular, we would like to thank the organizing committee for their tireless hard work. Finally, we would like to thank the Springer publications for producing this volume.

Dr. V. Bindhu
Conference Chair, ICCCES 2020
Head of the Department
Department of ECE, PPG Institute of Technology
Coimbatore, India

João Manuel R. S. Tavares
Porto, Portugal

Alexandros-Apostolos A. Boulogeorgos
Thessaloniki, Greece

Chandrasekar Vuppalapati
San Jose, USA

Contents

Security System for Big Data Cloud Computing Using Secure Dynamic Bit Standard 1
Faiz Akram

Distributed DBSCAN Protocol for Energy Saving in IoT Networks 11
Mazin Kadhum Hameed and Ali Kadhum Idrees

Hardware Implementation of Automatic Power Source Controller Cum Power Optimizer 25
A. Kunaraj, J. Joy Mathavan, and K. G. D. R. Jayasekara

Texture-Based Face Recognition Using Grasshopper Optimization Algorithm and Deep Convolutional Neural Network 43
Sachinkumar Veerashetty and Nagaraj B. Patil

An Interactive Framework to Compare Multi-criteria Optimization Algorithms: Preliminary Results on NSGA-II and MOPSO 61
David F. Dorado-Sevilla, Diego H. Peluffo-Ordóñez, Leandro L. Lorente-Leyva, Erick P. Herrera-Granda, and Israel D. Herrera-Granda

A Dynamic Programming Approach for Power Curtailment Decision Making on PV Systems 77
Yasmany Fernández-Fernández, Leandro L. Lorente-Leyva, Diego H. Peluffo-Ordóñez, and Elia N. Cabrera Álvarez

Arabic Braille Numeral Recognition Using Convolutional Neural Networks 87
Shurouq Alufaisan, Wafa Albur, Shaikha Alsedrah, and Ghazanfar Latif

In-Silico Modeling of Sleep Stage Classification System Using Covariance Normalization 103
Nishant Gautam, G. B. Mrudula, and C. Santhosh Kumar

Comparative Analysis of LSTM, One-Class SVM, and PCA to Monitor Real-Time Malware Threats Using System Call Sequences and Virtual Machine Introspection 113
Jayesh Soni, Suresh K. Peddoju, Nagarajan Prabakar, and Himanshu Upadhyay

SBP: Preparation of Schema Elements to Schema Matching 129
Aola Yousfi, Moulay Hafid El Yazidi, and Ahmed Zellou

Acceptance of Biometric Authentication Security Technology on Mobile Devices 145
William Ratjeana Malatji, Tranos Zuva, and Rene Van Eck

Impact of Agile Methodology Use on Project Success in South African Banking Sector 157
Makoena Moloto, Anneke Harmse, and Tranos Zuva

Automated Industrial Sound Power Alert System 175
R. Vishva, P. Harish Annamalai, K. Raja Raman, B. Vijay, J. Rolant Gini, and M. E. Harikumar

Predicting Early Readmission of Diabetic Patients: Toward Interpretable Models 185
Mir Moynuddin Ahmed Shibly, Tahmina Akter Tisha, and Md. Mahfuzul Islam Mazumder

Design and Analysis of Mobile-Based Tourism Security Application: Concepts, Artifacts and Challenges 201
Deepanjal Shrestha, Tan Wenan, Bikram Adhikari, Deepmala Shrestha, Adesh Khadka, and Seung Ryul Jeong

An Efficient Approach to Software Fault Prediction 221
Md. Anuvob Pradan, Mahia Binte Mizan, Moon Howlader, and Shamim Ripon

Indoor Navigation Assistant for Visually Impaired (INAVI) 239
U. B. Mahadevaswamy, D. Aashritha, Nikhil S. Joshi, K. N. Naina Gowda, and M. N. Syed Asif

On the Evaluation of Effectiveness of eLearning and Blended Learning 255
Sarka Hubackova

Design and Performance Analysis of GaAs-Based P-i-N Photovoltaic Using AlGaAs as Window Layer 263
Rocky Chakma, S. S. Mahtab, M. J. Alam, and Rupa Akter

Impact Analysis of Anti-Reflection Coating on P-i-N Solar Device 279
S. S. Mahtab, Rupa Akter, Tauhidul Mahbub, Ahamed Raihan, Rapsan Amin Anonto, and M. J. Alam

Microstrip Antenna with SVASTU Slot for UWB Applications 289
Shivani Chourasia, Sudhir Kumar Sharma, and Pankaj Goswami

Design and Implementation of a Smart Helmet System for Underground Miner's Safety 301
S. M. Minhajul Alam, Arnob Barua, Ahamed Raihan, M. J. Alam, Rocky Chakma, S. S. Mahtab, and Chitra Biswas

A Brief Review on Instance Selection Based on Condensed Nearest Neighbors for Data Classification Tasks 313
Yasmany Fernández-Fernández, Diego H. Peluffo-Ordóñez, Ana C. Umaquinga-Criollo, Leandro L. Lorente-Leyva, and Elia N. Cabrera-Alvarez

Virtual Group Movie Recommendation System Using Social Network Information 325
Tranos Zuva and Keneilwe Zuva

Hindi to English: Transformer-Based Neural Machine Translation 337
Kavit Gangar, Hardik Ruparel, and Shreyas Lele

CNN-Based Alphabet Identification and Sorting Robotic Arm 349
Saleh Al-Faraj, Mustafa Al-Bahrani, Saeed Al-Ghamdi, Marwan Rafie, Abul Bashar, and Ghazanfar Latif

Lung Cancer Detection from LDCT Images Using Deep Convolutional Neural Networks 363
Shahad Alghamdi, Mariam Alabkari, Fatima Aljishi, Ghazanfar Latif, and Abul Bashar

Global Integration and Distribution of Data Through Machine Learning for COVID-19 375
E. Laxmi Lydia, Jose Moses Gummadi, Chinmaya Ranjan Pattanaik, G. Jaya Suma, A. Krishna Mohan, and Ravuri Daniel

Role of Internet of Things and Machine Learning in Finding the Optimal Path for an Autonomous Mobile Robot 383
Dadi Ramesh, Mohmmad Sallauddin, Syed Nawaz Pasha, and G. Sunil

Overview of Blockchain Technology: Applications and Use Cases 393
J. S. Shyam Mohan, Vedantham Hanumath Sreeman, Vanam Venkata Chakradhar, Harsha Surya Abhishek Kota, Nagendra Panini Challa, M. U. M. Subramanyam, and Surekuchi Satya Swaroop

Tuberculosis Detection from CXR: An Approach Using Transfer Learning with Various CNN Architectures 407
P. Anu Priya and E. R. Vimina

E-commerce Logistic Route Optimization Deciphered Through Meta-Heuristic Algorithms by Solving TSP 419
M. Soumya Krishnan and E. R. Vimina

A Comparative Study on the Performance of Deep Learning Algorithms for Detecting the Sentiments Expressed in Modern Slangs 437
Vivank Sharma, Shobhit Srivastava, B. Valarmathi, and N. Srinivasa Gupta

Exploration of Varied Feature Descriptors for Diabetic Retinopathy Through Image Classification 449
K. Sreekumar and E. R. Vimina

News Topic Classification Using Machine Learning Techniques 461
Pramod Sunagar, Anita Kanavalli, Sushmitha S. Nayak, Shriya Raj Mahan, Saurabh Prasad, and Shiv Prasad

Blockchain for Management of Natural Resources Using Energy Trading as a Platform 475
S. Thejaswini and K. R. Ranjitha

Machine Learning Model for Anomaly Detection in Big Data for Health Care Applications 487
M. G. Sharavana Kumar and V. R. Sarma Dhulipala

Pilot Decontamination Algorithm with Iterative Weighted Graph Coloring Scheme for Multi-cell MIMO System in 5G Applications 501
V. Baranidharan, S. Karthikeyan, R. Hariharan, T. Mugunthan, and S. Vhivek

Synthesis Approach for Emotion Recognition from Cepstral and Pitch Coefficients Using Machine Learning 515
S. Vaijayanthi and J. Arunnehru

Performance Analysis of Deep CNN Assisted Optimized HIV-I Protease Cleavage Site Prediction with Hybridized Technique 529
Navneet Kaur and Wiqas Ghai

Artificial Neural Network for Identification and Classification of Natural Body Marks 541
Dayanand G. Savakar, Danesh Telsang, and Anil Kannur

Improvised Distributed Data Streaming Scheduler in Storm 557
J. Geetha, D. S. Jayalakshmi, Riya R. Ganiga, Shaguftha Zuveria Kottur, and Tallapalli Surabhi

Efficient Mining of Rare Itemsets 569
Shwetha Rai, Geetha M., and Preetham Kumar

EDoS-BARRICADE: A Cloud-Centric Approach to Detect, Segregate and Mitigate EDoS Attacks 579
S. B. Ribin Jones and N. Kumar

User Query-Based Automatic Text Summarization of Web Documents Using Ontology 593
K. Selvakumar and L. Sairamesh

Quantum Chaos-Based Encryption Technique for Transmission of Medical Images 601
R. Anitha and B. Vijayalakshmi

Secure Voting for Democratic Elections: A Blockchain-Based Approach 615
Hardik Ruparel, Shraddha Hosatti, Mahesh Shirole, and Sunil Bhirud

Convergence Analysis of Self-Adaptive Equalizers Using Evolutionary Programming (EP) and Least Mean Square (LMS) 629
N. Shwetha and Manoj Priyatham

Modified E-Shaped Resonator-Based Microstrip Dual-Mode Bandpass Filter 647
Shobha I. Hugar, Vaishali Mungurwadi, and J. S. Baligar

An Analysis of Air Quality by Vehicular Exhausts 653
V. Kanpurrani and A. L. Vallikannu

Secured Data Encryption and Decryption for Transmission in Device Independent Systems 663
B. N. Lakshmi Narayan, N. H. Prasad, and A. Mariyan Richard

Electrical Performance of Single Gate and Double Gate MOSFET to Optimize the Gate Length 673
James Lalthlamuana, Niladri Pratap Maity, L. L. K. Singh, and Reshmi Maity

Prediction of Diabetes Disease Using Machine Learning Model 683
Amandeep Sharma, Kalpna Guleria, and Nitin Goyal

Infrared Small Target Detection Based on Phase Fourier Spectrum Transform 693
Sur Singh Rawat, Sashi Kant Verma, and Yatindra Kumar

Knowledge Based Analytical Tool for Identifying Children with Dyscalculia 703
A. Devi and G. Kavya

RSA Algorithm Using Performance Analysis of Steganography Techniques in Network Security 713
B. Barani Sundaram, N. Kannaiya Raja, Nune Sreenivas, Manish Kumar Mishra, Balachandra Pattanaik, and P. Karthika

Performance of Two-Link Robotic Manipulator Estimated Through the Implementation of Self-Tuned Fuzzy PID Controller 721
Aditi Saxena, Rishabh Chaturvedi, and Jitendra Kumar

Rule Based Part of Speech Tagger for Arabic Question Answering System 733
Samah Ali Al-azani and C. Namrata Mahender

Change Detection in Land Use-Land Cover Using Convolutional Neural Network 745
Sahithi Samudrala, Mekala Pranvitha nand, Suhail Mohammad, and Radhesyam Vaddi

An Efficient Ant Colony-Based Self-adaptive Routing Strategy for IoT Network Model 753
K. M. Baalamurugan, D. Vinotha, and Premkumar Sivakumar

Optimizing Node Coverage and Lifespan of Wireless Body Area Network Using Hybrid Particle Swarm Optimization 765
S. Selvaraj and R. Rathipriya

Collective Examinations of Documents on COVID-19 Peril Factors Through NLP 779
E. Laxmi Lydia, Jose Moses Gummadi, Chinmaya Ranjan Pattanaik, B. Prasad, CH. Usha Kumari, and Ravuri Daniel

Design and Development of Topic Identification Using Latent Dirichlet Allocation 789
P. Lakshmi Prasanna, S. Sandeep, V. Kantha Rao, and B. Sekhar Babu

Hand Gesture Controlled Robot 799
N. H. Prasad, A. Mariyan Richard, and B. N. Lakshmi Narayan

Interdependence in Artificial Intelligence to Empower Worldwide COVID-19 Sensitivity 809
E. Laxmi Lydia, Jose Moses Gummadi, Chinmaya Ranjan Pattanaik, A. Krishna Mohan, G. Jaya Suma, and Ravuri Daniel

Design VLSI Architecture for 2_D DWT Using NEDA and KSA Technique 821
Satyendra Tripathi, Bharat Mishra, and Ashutosh Kumar Singh

Author Index 833

About the Editors

V. Bindhu received the B.E. degree in Electronics and Communication Engineering from Bharathiar University, Coimbatore, in 2002, M.E. degree in Applied Electronics from Anna University, Chennai, in 2007, and Ph.D. degree from Anna University, Chennai, in 2014. She has 10 years of teaching experience and 5 years of research experience. Currently, she is Professor at PPG Institute of Technology, Coimbatore. Her area of interest includes signal processing and VLSI design.

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 a 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 da Universidade do Porto (FEUP).

João Tavares is co-editor of more than 60 books, co-author of more than 50 book chapters, 650 articles in international and national journals and conferences, and 3 international and 3 national patents. He has been a 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 & 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 MSc and PhD 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.

Alexandros-Apostolos A. Boulogeorgos was born in Trikala, Greece, in 1988. He received the Electrical and Computer Engineering (ECE) diploma degree and Ph.D. degree in Wireless Communications from the Aristotle University of Thessaloniki (AUTh) in 2012 and 2016, respectively. From November 2012, he has been a member of the wireless communications system group (WCSG) of AUTh, while, from November 2017, he has joined the Department of Digital Systems, University of Piraeus. He also serves as an adjunct professor in the department of ECE in the University of Western Macedonia and as a visiting lecturer at the University of Thessaly. Finally, he is a Senior Member of the IEEE as well as the Technical Chamber of Greece, and serves as editor in IEEE Communications Letters.

Chandrasekar Vuppalapati is Software IT Executive with diverse experience in Software Technologies, Enterprise Software Architectures, Cloud Computing, Big Data Business Analytics, Internet Of Things (IoT), and Software Product & Program Management. Chandra held engineering and product leadership roles at GE Healthcare, Cisco Systems, Samsung, Deloitte, St. Jude Medical, Lucent Technologies, and Bell Laboratories Company. Chandra teaches Software Engineering, Mobile Computing, Cloud Technologies, and Web & Data Mining for Masters Program in San Jose State University. Additionally, Chandra held market research, strategy, and technology architecture advisory roles in Cisco Systems and Lam Research and performed Principal Investigator role for Valley School of Nursing where he connected Nursing Educators & Students with Virtual Reality technologies. Chandra has functioned as Chair in numerous technology and advanced computing conferences such as IEEE Oxford, UK, IEEE Big Data Services 2017, San Francisco USA, and Future of Information and Communication Conference 2018, Singapore. Chandra graduated from San Jose State University Masters Program, specializing in Software Engineering, and completed his Master of Business Administration from Santa Clara University, Santa Clara, California, USA.

Security System for Big Data Cloud Computing Using Secure Dynamic Bit Standard

Faiz Akram

Abstract Big data (BD) is a high-volume resource that requests savvy and imaginative types of data handling for improved knowledge revelation, dynamic, and procedure streamlining. CC gives a solid, deficiency open-minded, and adaptable condition to the enormous information appropriated the board frameworks. The Secure Dynamic Bit Standard (SDBS) calculation gives the security through two unique keys, for example, the ace key and meeting key created by cloud service providers (CSP). The SDBS calculation contains the three distinctive key lengths, for example, 128 bits, 256 bits, and 512 bits. The length of the ace key and meeting key is arbitrarily created during the encryption procedure. CSP scrambles the ace key with the meeting key and sends the encoded ace key and meeting key to the information suppliers on demand. The information supplier unscrambles the ace key with the meeting key and encodes the information record with the decoded ace key. This strategy will decrease the most extreme number of unauthenticated and unapproved clients in a huge information cloud.

Keywords Cloud computing (CC) · Cloud service provider (CSP) · Secure Dynamic Bit Standard (SDBS)

1 Introduction

Cloud computing (CC) has numerous applications, for example, empowering access to costly applications at no cost, lessening both foundation and running costs of PCs and programming as there is no requirement for any establishment. Clients can put the information at anyplace [1]. All clients are required to connect with a framework, state the Internet. CC began as an instrument for relational figuring, however now it is generally utilized for getting to programming on the web, online capacity without stressing over foundation cost, and preparing power. Associations can offload their information technology (IT) foundation in the cloud and get entrance. Not just private

---

F. Akram (✉)
Faculty of Computing and Informatics, Jimma Institute of Technology, Jimma University, Jimma, Ethiopia
e-mail: akram.faiz@ju.edu.et; akram.faiz@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_1

associations are moving to distributed computing, yet the legislature is additionally moving a few pieces of its IT framework to the cloud. Huge information includes the advanced information from a few computerized sources which contain sensors, scanners, numerical displays, recordings and cell phones, digitizers, Internet, messages, and interpersonal organizations which are expanding the information rate.

CC and enormous information are conjoined [2, 3]. BD gives clients the capacity to utilize merchandise figuring to process dispersed questions over numerous datasets and return resultant sets in an opportune way. CC gives a class of dispersed information handling stages. Huge information sources from the cloud and Web are put away in a conveyed flaw lenient database and handled through a programming model for huge datasets with an equal disseminated calculation in a group [4]. The multifaceted nature and assortment of information types are handling the capacity to perform an investigation on huge datasets. CC foundation can fill in as a successful stage to address the information stockpiling required to perform an enormous information examination. CC is associated with another example for the arrangement of registering foundation and enormous information handling strategy for a wide range of assets accessible in the cloud through information examination.

A few cloud-based advancements need to adapt to this new condition since managing BD for simultaneous handling has gotten progressively confounded. MapReduce is a genuine case of BD handling in a cloud domain; it permits the preparation of a lot of datasets put away in equal in the group. Group processing displays great execution in conveyed framework situations, for example, PC force, stockpiling, and system correspondences. Moreover, the capacity of bunch figuring is to give a cordial setting to information development. Database management systems (DBMSs) are viewed as a piece of the currently distributed computing engineering and assume a significant job to guarantee the simple change of utilizations from old venture foundations to new cloud framework designs [5]. The weight for associations to rapidly receive and execute advances, for example, CC, to address the test of large information stockpiling and preparing requests involves unforeseen dangers and results [6].

2 Security in BD-CC

Big data cloud computing (BD-CC) transforms into a valuable and standard plan of action in light of its engaging segments. Notwithstanding the current advantages, the past segments also bring about authentic cloud-specific security issues. The all-inclusive community concern is security in the cloud, and the clients are deferring to trade their business to the cloud. Security issues have been the prevention of the improvement and expansive use of CC [7]. Understanding the security and assurance chances in CC, making rich and powerful arrangements are essential for its thriving, regardless of the way that mists empower clients to avoid fire up costs, lessen working expenses, and unite their speed by rapidly getting administrations and infrastructural assets when required.

In publicizing and business, the greater part of the ventures utilizes large information; however, the key properties of security may not be executed [8]. On the off chance that security penetrates happens to BD, it would result in much more genuine lawful repercussions and reputational harm than at present.

In this new time, numerous organizations are utilizing the innovation to store and examine petabytes of information about their organization, business, and their clients. For making BD secure, procedures, for example, encryption, logging, nectar pot recognition must be fundamental. In numerous associations, the organization of BD for misrepresentation location is alluring and helpful. The test of recognizing and forestalling propelled dangers and noxious interlopers must be explained utilizing large information style examination. The difficulties of security in CC conditions can be classified into an organized level, client verification level, information level, and conventional issues [9].

The difficulties can be sorted under a system-level arrangement with organizing conventions and system security, for example, disseminated hubs, and appropriated information, and inter hub correspondence.

The difficulties can be classified under client verification level arrangements with encryption or unscrambling procedures, validation techniques, for example, authoritative rights for hubs, verification of utilizations and hubs, and logging.

The difficulties can be classified under information level arrangements with information honesty and accessibility, for example, information security and disseminated information.

The difficulties that can be ordered under the general level are customary security apparatuses and utilization of various advancements.

3 Security Attacks

Distributed computing relies for the most part upon the structure of the current system, for example, metropolitan area network (MAN), wide area network (WAN), and local area network (LAN). System-level security assaults might be deliberately made by outside clients or by a malevolent insider staying between the client and the cloud service providers (CSP) and endeavoring to encroach upon the information to or from the cloud. This segment will endeavor to focus on the system-level security assaults and their possible countermeasures to ensure genuine information secrecy and dependability [10, 11].

3.1 Space Name System (DNS)

Attacks on the Internet, since reviewing a framework described by the numbers is troublesome, the aggressors are made to do with names. The Internet Protocol plays a remarkable role over the Internet related to the PC. The names of the DNS automatically change concerning the IP addresses utilizing a disseminated database schema. Web DNS servers are dependent upon various types of attacks, for example, space catching, ARP store hurting, and man-in-the-middle attacks. A discussion of these attacks was found underneath.

3.2 Space Hijacking

Domain capturing alludes to changing the name of an area without the data or consent from the area owner or producer. Area seizing engages intruders to get corporate data and play out the unlawful development, for instance, phishing, where a site is superseded by a similar segment that records private data. Another arrangement is using the Extensible Provisioning Protocol (EPP) that is used by various area vaults. EPP utilizes an approval code gave uniquely to the space registrant as a safety effort to envision unapproved names advancing.

3.3 IP Spoofing Attack

The attacker buildups across unapproved admittance to a PC by proposing the PC to have the intention of undergoing through heavy traffic are known as IP mocking. Various attacks are employed by IP caricaturing, for instance, denial of service assault [12].

3.4 Disavowal of Service (DoS) Attacks

The inspiration driving these assaults is making the physical system and PC assets out of reach. During the DoS attack, the attacker submerges the disaster with a broad number of software packages over short intervals. DoS aids to identify distinguishing the devouring data between the systems. The assailant practices a false IP address as the source IP address to hold the interruption of the DoS. Furthermore, it is conceivable to the aggressor to utilize distinctive traded off machines that need to begin at presently seized to attack the misfortune machine in the interim. This kind of interruption handled by the DoS is known as the distributed DoS [13].

3.5 Transmission Control Protocol Synchronize (TCP SYN) Flooding Attack

While possessing situations like the DoS attack, the hackers employ the TCP SYN packages to pervert the machines. These kinds of attacks will certainly damage the restrictions of the three-course handshake when maintaining the half-open affiliations. A man-in-the-middle (MITM) attack is an overall term for when a culprit positions himself in a discussion between a client and an application by either to snoop or to imitate one of the gatherings, causing it to show up as though an ordinary trade of data is in progress. The bundle filtering is a method adopted to reduce the IP ridiculing that has been executed with the assistance of beginning stage confirmation systems, solid encryption, and a firewall [14, 15].

4 Methodology

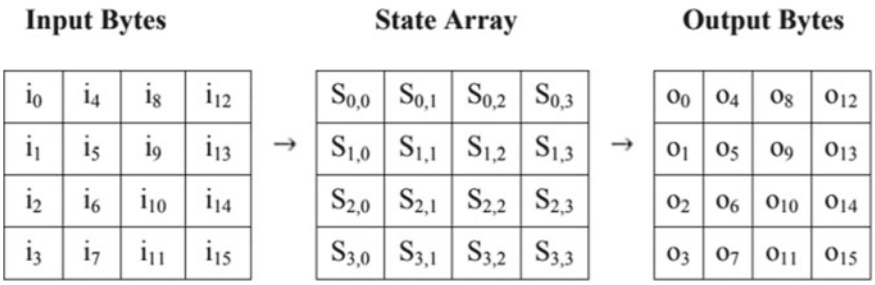
SDBS have three-piece levels, for example, 128 bits, 256 bits, and 512 bits. At whatever point the information supplier needs to transfer an information record to the cloud, any of the bit levels is chosen haphazardly and it will get changed over into bytes. Because of byte esteem, CSP will produce the ace key and a meeting key utilizing an irregular generator. The ace key is scrambled by the meeting key, and both the encoded ace key and the meeting key will be sent to the information supplier.

The meeting key will be utilized to unscramble the ace key, and the ace key is utilized to scramble the information document. The encoded information document will be transferred to the huge information cloud server alongside the proof of ownership which is created by CSP. On the off chance that an information client needs to download an information document from the large information cloud server, according to popular demand, the CSP will send the encoded ace key and the meeting key alongside the scrambled information record to the information client after one-time password (OTP) check.

At that point, utilizing this meeting key the ace key will be unscrambled, and utilizing the ace key the information record will be decoded and put away in the arrangement of the information client. SDBS calculation is a novel calculation that has three different guidelines with ten rounds for 256-bit keys, eight rounds for 128-bit keys, and 12 rounds for 512-bit keys. The round contains different activities like substitution, adjustment, and change of the info plaintext into the yield figure text (Fig. 1).

The encryption process of SDBS algorithm with the 128-bit standard is shown in Fig. 2.

The plain content called a state cluster ‘S’ has four lines of bytes. Each line of a state contains N_b quantities of bytes, where N_b fluctuates for these three guidelines.



i_0	i_4	i_8	i_{12}
i_1	i_5	i_9	i_{13}
i_2	i_6	i_{10}	i_{14}
i_3	i_7	i_{11}	i_{15}

$S_{0,0}$	$S_{0,1}$	$S_{0,2}$	$S_{0,3}$
$S_{1,0}$	$S_{1,1}$	$S_{1,2}$	$S_{1,3}$
$S_{2,0}$	$S_{2,1}$	$S_{2,2}$	$S_{2,3}$
$S_{3,0}$	$S_{3,1}$	$S_{3,2}$	$S_{3,3}$

o_0	o_4	o_8	o_{12}
o_1	o_5	o_9	o_{13}
o_2	o_6	o_{10}	o_{14}
o_3	o_7	o_{11}	o_{15}

Fig. 1 Input bytes, state array, and output bytes

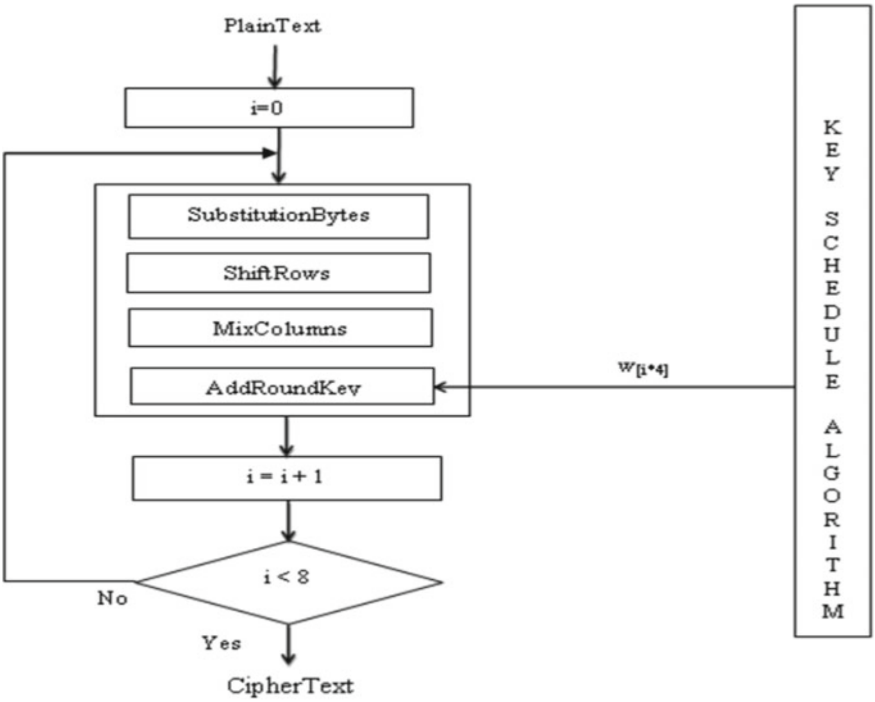


Fig. 2 SDBS encryption process for 128-bits standard

For the 128-bit standard, the estimation of N_b is 4, for the 256-bit standard, the estimation of N_b is 8, and for the 512-bit level, the estimation of N_b is 16.

The variety of information bytes appeared as $i_0, i_1, \ldots, i_{15}$ and the variety of yield bytes is spoken to by $o_0, o_1, \ldots, o_{15}$ as appeared in Fig. 1.

Decoding process: In SDBS unscrambling process, the information supplier must login and afterward select an information record, and the information client needs to download. Before downloading the information document, the information client

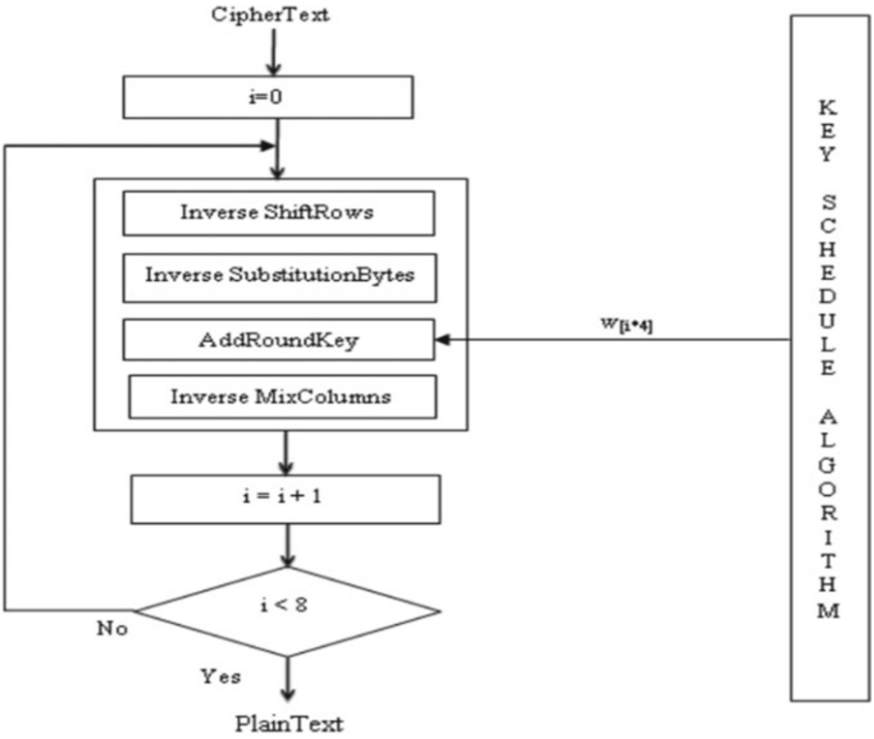


Fig. 3 SDBS decryption process of 128-bits standard

sends the solicitation to the CSP to get the ace key and meeting key. CSP encodes the ace key with the meeting key and sends to the information client.

OTP is additionally sent by the CSP to the information client's email id or versatile no. On the off chance that the entered OTP is substantial, at that point the encoded information document will be downloaded from the cloud, and the decoded ace key with the meeting key is utilized to begin the unscrambling procedure. At long last, the decoded record is put away into the framework (Fig. 3).

5 Results

The proposed method is performed and tried on a workstation with the following hardware specifications like 8 GB RAM, Windows 7 (64-bit) operating system, Intel i7 processor within the cloud storage (dropbox distributed storage) (Table 1).

The exhibition of SDBS calculation could be dissected by two sorts of boundaries which are encryption time and decoding time. The encryption time is measured by the time taken to complete the encryption process along with the record size.

Table 1 Role and operation of multilevel SDBS

Role	Operation
Data Provider (DP)	Encrypt Data File
	Upload Data File
Data User (DU)	Download Data File
	Decrypt Data File
Cloud Service Provider (CSP)	Authenticate—Checking User Name and Password
	Authorize—Checking Credentials of the User
	Key Generation, OTP Generation, PoW Generation
	Block Unauthorized User

And the decryption time is measured by the time taken to complete the decryption process along with the record size. The genuine portrayal of SDBS 128-bits standard encryption time and unscrambling time-dependent on record size is spoken to underneath.

Figure 4 portrays the presentation of encryption time and unscrambling time versus record size in conspire 4. The diagram is completely based on the encryption time and document size. The encryption or unscrambling time of the proposed scheme 4 is less time when contrasted and different methods conspire 1, plot 2, and plan 3. This strategy took 2.56 ms for encoding 1 GB information document, likewise the 24 GB information record took 12.1 ms for encryption. From Fig. 4, 15, 1 GB information record took 2.35 ms for decoding; additionally, the 24 GB information document took 11.72 ms for unscrambling.

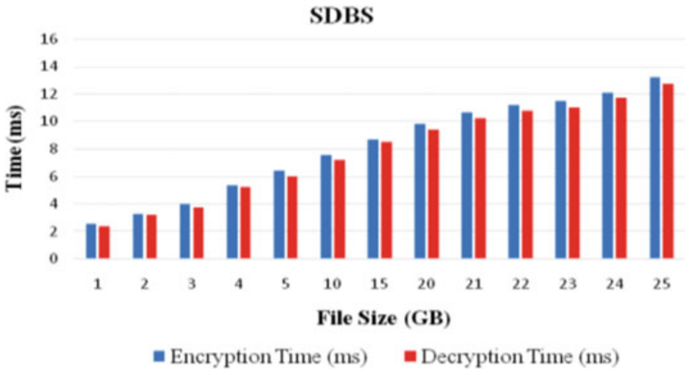


Fig. 4 Performance of SDBS 128-bits standard

6 Conclusion

The proposition gave a point by point prolog to the security framework in big data cloud computing. Big data distributed storage limits and applications are clarified. The conceivable outcomes of various assaults are portrayed in detail. In this article, information moved to the big data cloud has been finished by the information supplier. The proposed framework has high information respectability and information stockpiling without information misfortune. Before the information has been transferred into the capacity region, a high secure calculation called Secure Dynamic Bit Standard is been utilized. Big data investigation report assists by distinguishing each datum supplier and information client use of document size, encryption time or decoding time, and transfer time or download time.

References

1. Nahar AK, Mongia K, Kumari S (2018) Cloud computing and cloud security. Int J Res Adv Eng Technol 4(1):1–8
2. AshwinDhivakar MR, Ravichandran D, Dakha V (2015) Security and data compression in cloud computing using BlobSeer technique. In: National conference on cloud computing and big data, vol 1(12), pp 201–203
3. Ateniese G, Fu K, Green M, Hohenberger S (2006) Improved proxy re-encryption schemes with applications to secure distributed storage. ACM Trans Inf Syst Secur 9(1):1–30
4. Attrapadung N, Herranz J, Laguillaumie F, Libert B, De Panafieu E, Ràfols C (2012) Attribute-based encryption schemes with constant size cipher texts. Theoret Comput Sci 422:15–38
5. Awodele O, Izang AA, Kuyoro SO, Osisanwo FY (2016) Big data and cloud computing issues. Int J Comput Appl 133(12):35–47
6. Bachhav S, Chaudhari C, Shinde N, Kaloge P (2016) Secure multi cloud data sharing using key aggregate cryptosystem for scalable data sharing. Int J Comput Sci Inf Technol 3(1):19–27
7. Balasubramanian N, Balasubramanian A, Venkataramani A (2009) Energy consumption in mobile phones: a measurement study and implications for network applications. In: Proceedings of the 9th ACM SIGCOMM conference on internet measurement conference, vol 1(5), pp 280–293
8. Bavisi S (2018) Computer and information security handbook. Morgan Kaufmann Publication, Elsevier Inc., pp 375–341
9. Bhadauria R, Chaki R, Chaki N, Sanyal S (2011) A Survey on security issues in cloud computing. IEEE Commun Surv Tutor 3(16):1–15
10. Bisong A, Rahman M (2011) An overview of the security concerns in enterprise cloud computing. Int J Netw Secur Appl 3(1):30–45
11. Hemalatha M (2012) Cloud computing for academic environment. Int J InfCommunTechnol Res 97–101
12. Mathew S (2012) Implementation of cloud computing in education—a revolution. Int J Comput Theory Eng 473–475
13. Kaur M, Singh H (2015) A review of cloud computing security issues. Int J AdvEngTechnol (IJAET), pp 397–403
14. Gaikwad BP (2014) A critical review on risk of cloud computing in commercial. Int J Innov Res Comput CommunEng 1–8
15. Ahmed ES, Saeed R (2014) A survey of big data cloud computing security. Int J Comput Sci SoftwEng (IJCSSE), pp 78–85

Distributed DBSCAN Protocol for Energy Saving in IoT Networks

Mazin Kadhum Hameed and Ali Kadhum Idrees

Abstract Sensor networks form a crucial topic in research, as it seems to target a huge variety of uses in which it could be applied, such as health care, smart cities, environment monitoring, military, industrial automation, and smart grids. The clustering algorithms represent an essential factor in conserving power within energy-constrained networks. The selection of a cluster head balances the energy load within the network in a proper way, eventually contributing to the reduction of energy consumed, as well as the enhancement of network lifespan. This article introduced a distributed DBSCAN protocol for saving the energy of sensor devices in IoT networks. This protocol is implemented on each IoT sensor device, and the devices apply the density-based spatial clustering of applications with noise (DBSCAN) algorithm to partition the network into clusters in a distributed way. The efficient periodic cluster head strategy is proposed based on certain criteria like remaining energy, number of neighbors, and the distance for each node in the cluster. The cluster head will be chosen in a periodic and distributed way to consume the power in a balanced way in the IoT sensor devices inside each cluster. The comparison results confirm that our protocol can conserve power and enhance the power conservation of the network better than other approaches.

Keywords Sensor networks · Density-based spatial clustering of applications with noise (DBSCAN) clustering · Wireless sensor network (WSN) · Low-energy adaptive clustering hierarchy (LEACH) · Internet of Things (IoT)

M. K. Hameed
Department of Software, University of Babylon, Babylon, Iraq
e-mail: it.mazen.kadhum@uobabylon.edu.iq

A. K. Idrees (✉)
Department of Computer Science, University of Babylon, Babylon, Iraq
e-mail: ali.idrees@uobabylon.edu.iq

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_2

1 Introduction

Wireless sensor networks (WSNs) have recently gained significant attention for their implications found in various fields including ecosystem monitoring, health care, environment assessment, urban areas applications, control maintenance, and target tracking [1]. The connection of all things that the Internet can monitor or control could be defined as Internet of Things (IoT), which is most preferably achieved through a wireless medium [2, 3]. The network of wireless sensors could be depicted as the set of huge sensor nodes used over a wide area for sensing and accumulating different data from the systems and environment, to be applied in a variety of uses like weather monitors, animal tracking, disaster managing, and bio-medical applications, within IoT [4, 5]. Wireless sensors could be of use with IoT applications for gathering and processing data with the extraction of valuable information to be communicated to the end user, as it can occasionally be unreachable by individuals. Therefore, WSNs are among the integrated parts of IoT applications [6]. Direct communication is made by each node over the BS as the data is distributed [4]. Through the continuity of data transmitted, the furthest node would be more likely to die earlier than others through its energy loss [7]. Consequently, the clustering process applied tends to collect nodes, forming a set of clusters for solving the problem [2]. Its main performance is remarkably improved through clustering several nodes [8]. Besides, the network keeps the demands of the central organization to a minimum, inspiring the local decision-making to enhance the scalability. The clustering procedure tends to collect data through the active network. As well, a suitable cluster head needs to be selected for every data retrieving clusters [9], through sensor nodes to be passed to the BS [10]. The SN eventually creates clusters for the monitoring procedure and the constitution of both cluster member (CM) and cluster head (CH) [7]. The SNs contain battery sources for their performance, which makes them a tool of power starving. The main elements that influence the energy dissipating of WNS are the distance with the sink, the remaining energy of the node, and intra-cluster distance [11]. The remained power, distance among the core nodes in the cluster, and the number of members in each core node are the three factors that the CH choosing is based on in our work. Thus, at the same time, the elected CH must have the highest remained power, maximum members, and finally the lowest distance to all core nodes. The power can be preserved to extend the life of the sensor device if these factors considering, and the simulations show the good results.

The contributions of the proposed work are based on the distributed clustering-based DBSCAN protocol which is for increasing the lifespan of wireless sensors of IoT networks. This protocol is distributed on every IoT sensor device, and the sensor devices are combined with the DBSCAN algorithm to form several groups in the network area of interest. An efficient periodic CH approach is suggested based on several criteria like remaining power, neighbors' number, and the distance for each node in the group. The cluster head will be elected periodically to the consumed power which is balanced in the wireless sensors inside each group. The proposed protocol is evaluated and compared to two existing methods such as I-LEACH [11] and low-energy adaptive clustering hierarchy (LEACH) that are presented in [12] in light of several metrics like the resting energy, network lifetime, and CH count, etc. The comparison of simulation results illustrates that our protocol can preserve energy and increase the network lifetime better than other approaches.

The remaining of the article is structured as follows. The related literature is presented in Sect. 2. The DBSCAN traditional algorithm is explained in more detail in Sect. 3. Section 4 introduces the proposed energy-efficient distributed clustering-based DBSCAN protocol for conserving the power of wireless sensors of IoT networks. Results and discussion are explained in Sect. 5. Section 6 presents the conclusion and the planned work for the future.

2 Related Work

The cluster head (CH) has been selected randomly resulting in a similar likeliness for both nodes with high or low energy to become a CH. Whenever a low energy node is elected to be a CH, it is most likely to die soon, with its eventual effect on how robust the network is. Besides, every round varies in how many CHs it has, as well as their location [13]. As with WSNs, the main focus lies on two essential factors: reducing the consumed power, as well as extending the network lifespan. Taking LEACH protocol to be the basic algorithm, several alterations are made in light of differing applications. LEACH and its related researches have been presented [14], taking into account several significant parameters including the clustering method, data aggregating, scalability, and mobility type. This protocol makes a random selection of CHs without the BS knowing any details on the network's remaining energy [15]. Therefore, the LEACH-C protocol has been suggested for addressing this issue [12].

The PSO-ECHSs are dealt with as the Particle Swarm Optimization (PSO) based CHs are selected with the use of factors such as distance among nodes, remaining energy, and distance to BS [16]. An alternative optimizing method known as grouped gray wolf search optimization has been applied in [17] selecting security-aware CH, to improve the network lifespan choice. The researchers proposed an alternative algorithm which first calculates the ideal cluster number, taking into account its location adaptability and data aggregating rate. Next, a new parameter is presented in light of the residual, initial, and average energy consumption, as well as the node degree for selecting the CH. A third aspect is the proposal of an unevenly self-adaptive clustering algorithm that considers the node degree in solving the "hot spot" problem. At last, a solution is suggested for the "isolated nodes problem" [18].

Jan et al. [19] present a new method known as a mutual authentication approach based on payload, as it consists of two steps selecting nodes optimality which act as CH, being allowed to have communication with its neighbors, and the authentication of every CH of its near nodes for forming clusters. After the former step, a method of authenticating takes place which depends on tokens. The tokens are used in the correlation between the CH and the acknowledge messages it corresponds with. Authenticating the payload cluster contributes for forming clusters from close member nodes and the CH. A comparison between the scheme of each the LEACH-C and Sec LEACH is suggested. The suggested model shows a lack in the use of an encryption method, as well as the improvement of its performance and comparing it to modern clustering models with a random distribution. Purohit and Bhargava introduced the multi-hop routing scheme [20], where the one-hop transmission is transformed into multi-hop way, to reduce the consumed energy by a sensor. This helps in obtaining the effective use of energy. The experiment resulted in the improvement of performance regarding the time for the first node to die (FND). This had a clear influence on improving network energy effectively. The main aim of this proposed idea is inter-cluster communicating; the time needed to receive messages is rather higher, negatively affecting its work. The main use of WSNs is receiving information for doing several performances, mainly in light of the data received. The decrease of messages counts automatically displays the inactive nodes, eventually declining the network's general performance. The work in [11] is introduced an improved method of the LEACH named I-LEACH. It is limited the selection of the cluster head using a certain threshold with concurrently changing the level of power between the nodes. The results explain better performance with the original protocol.

Despite many clustering approaches were proposed for grouping the sensor nodes in the WSNs, but none of them can ensure an optimal energy saving, and this would result in a shorter lifespan of the WSN. This paper suggested a distributed DBSCAN protocol for preserving the power of sensor devices in IoT networks. This protocol is executed at each IoT wireless sensor, and the devices apply the DBSCAN scheme to divide the network into groups in a distributed way. An efficient cluster head strategy is proposed based on certain criteria like remaining energy, number of neighbors, and the distance for each node in the cluster. The cluster head will be selected in a periodic and distributed way to conserve the power in a balanced way in the IoT wireless sensors inside each cluster.

3 The DBSCAN Algorithm

The DBSCAN approach identifies clusters within huge spatial datasets by taking into account the local density of its elements, with the use of a single input parameter. Also, a suitable parameter value is suggested for the user, so little knowledge about the domain itself is needed. The aim of DBSCAN is categorizing them into clusters apart, eventually defining the differing classifications [21].

In traditional DBSCAN, the user requires two parameters to be defined that are the neighborhood range and MinPts refer to the lower required number of points to construct a new cluster. In the beginning, the scheme chooses one-point P randomly and then calculate the distance among this selected point and the rest point in the dataset. The neighborhood condition between the point P and any other point in the dataset if the distance between them is less or equal to e. If the number of points which are in the neighborhood range of point P greater than or equal to MinPts, then the new cluster will be constructed; otherwise, these points as noise are labeled.

This means that noise points can later be within other clusters if they satisfy the condition of the required MinPts in the neighborhood range of the newly selected point. After that, the DBSCAN scheme will check if it is possible to extend this cluster or it chooses another point from the outside of the current cluster. The checking is done by verifying both MinPts and distance conditions if they are satisfied for each point in the range of the cluster. If these conditions are satisfied, then the DBSCAN scheme extends this cluster to each point in the neighborhood range of point P. The extension of the cluster will be stopped, and each point will be labeled as a visited point if the cluster expanded to the required the MinPts. The DBSCAN scheme then chooses another random not visited point from the dataset and repeats the same above scheme. The DBSCAN scheme will be stopped if there is no point labeled as not visited. Algorithm 1 illustrates the DBSCAN scheme with its expanding function. The time requirement of the DBSCAN scheme is O(n2), where n refers to the size of the dataset. The time complexity will be decreased to the O(nlogn) if the spatial indexing is utilized [22, 23].

4 Energy-Efficient Distributed DBSCAN Protocol

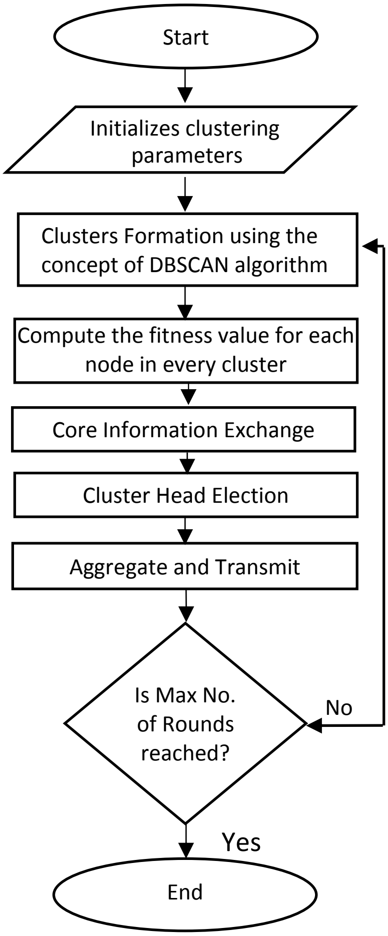
This research aims at proposing a distributed DBSCAN protocol for maximizing the wireless sensor's lifetime. This protocol is distributed at every sensor device deployed in the monitored area. The proposed protocol involving two steps: setup and steady-state. For the sake of simplicity, the proposed distributed DBSCAN protocol is named as DBSCAN protocol in this paper. The DBSCAN protocol is presented in Fig. 1.

In the setup phase, when the sensor devices are deployed in the working area, it is supposed that every sensor device knows its location. According to the DBSCAN algorithm, each sensor implements the algorithm of DBSCAN as follows:

(a) Each sensor node will perform the same test whether it is core point or not, according to the principle of the DBSCAN algorithm, as it scans the surrounding area to find out the number of sensors that are within the sensing range and that must be greater or equal to a specific parameter.
(b) As for the sensor nodes that are within the sensing range for core point, it will be its member.
(c) If this core point does not belong to any cluster, then it forms a new cluster; otherwise, it remains with the same cluster.
(d) The core point sends a message to all its members to be included in the same cluster.
(e) Repeat steps 1–5 until all sensor nodes are passed.

Algorithm 1 explains the distributed DBSCAN algorithm that will be executed in every sensor node *sj*.

Fig. 1 DBSCAN protocol


Start
Initializes clustering parameters
Clusters Formation using the concept of DBSCAN algorithm
Compute the fitness value for each node in every cluster
Core Information Exchange
Cluster Head Election
Aggregate and Transmit
Is Max No. of Rounds reached?
No
Yes
End

Algorithm 1. Distributed DBSCAN (sj)

Input: N: number of neighbor nodes, Sr: sensing range, minNodes: minimum number of nodes to create cluster.

Output: sj.rejon: the cluster number for node sj.

```
while  $REj \geq Ethr$ do
    If  sj  Receive MemberPacket from si  then
        Mark sj  as member to the Core si ;
      Update REj;
end
      sj.rejon $\leftarrow$ 0;
      for each node si in N do  // $i \in N$ and $i \neq j$
      nbrNodes $\leftarrow$ nbrNodes + CORE Objective Function (sj, si, Sr);
      if  CORE Objective Function return 1 then
                  Send MemberPacket to the sensor node i;
                  Update REj;
      end
      if  nbrNodes $\geq$ minNodes then
              save the information
              if (((sj.rejon = 0) Or (sj.rejon $\neq$ 0)) and (r==0)) then
                sj.rejon $\leftarrow$ sj.rejon +1;
                   Call Cluster(sj);
               end
               else if ((sj.rejon = 0) Or (r $\neq$ 0)) then
                       sj.rejon $\leftarrow$ sj.rejon +1;
                          Call Cluster1(sj);
                end
                 else if ((sj.rejon $\neq$ 0) Or (r $\neq$ 0)) then
                            Call Cluster2(sj);
                 end
               end
        end for
    end while
returen sj.rejon;
```

CORE objective function return 1 and $r = 0$ if the sensor node *i* is within the sensing range Sr, and it is not a member in other clusters. Otherwise, CORE objective function returns 0 and $r = 1$. The function cluster put any neighbor node within the sensing range of *sj* in the same cluster and *sj* send MemberPacket to the sensor node *i* to 4 cluster of *sj*. The function cluster1 places any neighbor node within the sensing range of *sj*, and it has not been assigned to any cluster in the same cluster of sensor node *j*. The function cluster2 places any neighbor node within the sensing range of *sj*, and it has not been assigned to any cluster (or it is assigned to the cluster of sensor node *j*) in the same cluster of sensor node *j*. After achieving the functions cluster, cluster1, and cluster2, the remaining energy of the sensor node *j* will be updated due to sending a MemberPacket to the sensor node *i* to inform it that it becomes a member in the same cluster of *sj*.

After the stage of creating clusters, the exchange of information between the core points (nodes) is done inside the single cluster, where each core point sends a message to all the core points inside the cluster; it contains all the necessary information inside like remained power, status, location, number of members, total wireless sensors number in the group (cluster), etc. Every core sensor node inside each cluster will include the information of other core nodes in the same cluster; therefore, every core node in the same cluster will execute Eq. (1) for the information of each member inside the core node. The core node that gives the better value of Eq. (1) will be selected as a cluster head in the current cluster for this round. All the core nodes inside the cluster will achieve the same computation and will give the same results for the winner core node. This will be implemented in a distributed way, and every core node will know if it is a cluster head or not.

$$\begin{aligned}\mathrm{FitVal}_j = {} & \frac{E_{\mathrm{remaining}}}{E_{\mathrm{initial}}} + \left(1 - \sum_{j \in N} \left| S_j(x, y) - S_i(x, y) \right| \right) \\ & + \frac{S_j(\text{ Members})}{\text{Cluster(Members)}} \end{aligned} \tag{1}$$

where $E_{\mathrm{remaining}}$ is the remaining power of the wireless node j; E_{initial} is the initial energy value of sensor node j; N is the number of core nodes in the current cluster; S_j (x, y) and S_i (x, y) refer to the positions of core nodes S_j and S_i, respectively. S_j (Members) refers to the number of nodes member of core node j; cluster (Members) refers to the total number of nodes in the cluster.

In the steady phase, after clusters formation and fixing the TDMA schedule, the process of transmitting data may start. With the assumption that all nodes contain data that requires to be sent to the CH, this sending will occur within its allocated time. The cluster head in its turn allocates a TDMA schedule to the actively participating members so that the data transmitting process is managed, and the consumed power is limited to a minimum. Based on the active/idle status, the data is transmitted during the steady-state stage with regards to the timespan assigned for every member node. The power supply of any idle node will be turned off but the CH, which awaits the data from the member nodes. After delivering data through nodes, the CH initiates a data aggregating process followed by transmitting this data to the BS. The algorithm eventually returns to the setup iterating stage to select a different CHs group, followed by the steady stage, and so on.

5 Performance Evaluation, Analysis, and Discussion

This section focuses on evaluating the proposed distributed DBSCAN protocol using different performance metrics like cluster count, remaining energy, dead nodes number, packets number transmitted to cluster head, number of packets sent to cluster sink, and network lifetime. The conduction of the simulation results is performed

Table 1 Stimulation parameters

Symbol	Description	Value
$X1_m$	Distance at X-axis	400 m
$Y1_m$	Distance at Y-axis	400 m
Ns	WSN size	100 nodes
$P_{\mathrm{Tx}}, P_{\mathrm{Rx}}$	The initial energy	0.5 J
E_{mp}	Energy consumption for receiving	0.0013/pJ/bit/m^4
E_{fs}	Energy dispersion: free space model	10/pJ/bit/m^2
E_{amp}	Energy dispersion: power amplifier	100/pJ/bit/m^2
E_{DA}	Energy consumption for collection	5/nJ/bit
d_0	Reference distance	87 m
I	Packet size	4000 bits

using a C++ custom simulator for 2500 iterating round so that several plots are obtained. The sensor nodes are deployed in the monitored area randomly. The location of the sink is in the center of the monitored area; with no limitation on energy, normal nodes would have its limitations in terms of energy, memory, and processing capabilities. The suggested protocol is applied for generating the results with regards to the parameters referred to in Table 1.

As for this work, the packet size tends to be relatively larger, namely be 4000 bits. The proposed protocol in this paper is named as DBSCAN. The DBSCAN protocol is applied to the same energy consumption model that is employed in [11]. An obvious result of stimulation is that the DBSCAN outperforms the I-LEACH [11] and LEACH [12] protocols in light of several performance metrics. The number of CHs is found to extend up to 2000 rounds for the DBSCAN, whereas it reaches only 1750 and 850 rounds for I-LEACH and LEACH, respectively. Figure 2 shows the cluster count for the LEACH, I-LEACH, and DBSCAN.

Similarly, the simulation results in Fig. 3a present the fact that at just "500 rounds," the average power of LEACH reaches 0, while I-LEACH goes on to ~"1250 rounds." The DBSCAN continues to 1750. The amount of data packets transmitted to BS within LEACH and I-LEACH reaches maximally (0.5 × 104) and (1.75 × 104), respectively. The rise in rounds leads to the energy depletion of sensor nodes until they terminate.

Figure 4 illustrates the network lifetime through the representation of dead nodes. After 750 rounds, the number of nodes alive levels out at 0 for LEACH, whereas a few nodes remain active till 1500 rounds with I-LEACH. The DBSCAN, on the other hand, has several nodes that remain active till 2000 rounds.

As for the DBSCAN, the value may reach (2×10^4), as is shown in Fig. 5a. A similar increase in data packets transmitted to cluster head (CH) is noticed in Fig. 5b, proving the effectiveness of the suggested protocol. This illustrates the

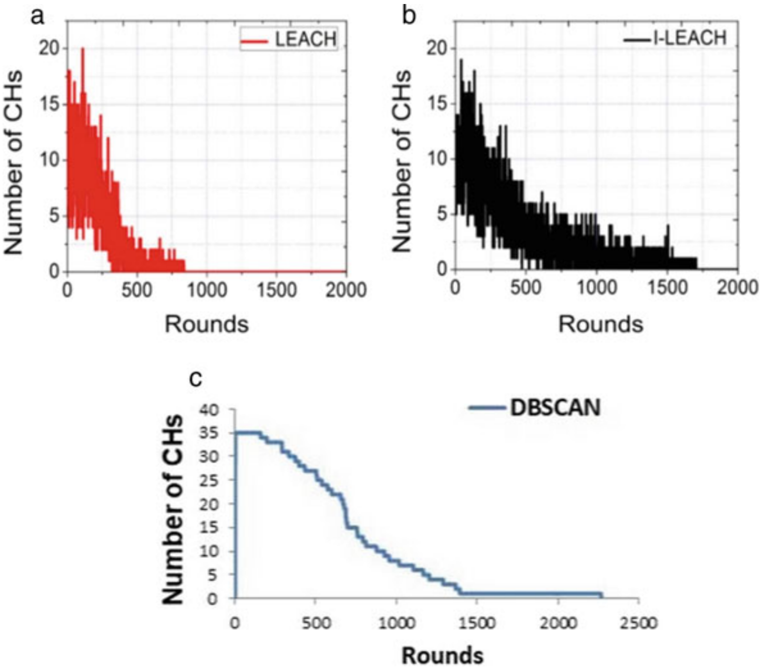


Fig. 2 CH count **a** LEACH, **b** I-LEACH, **c** DBSCAN

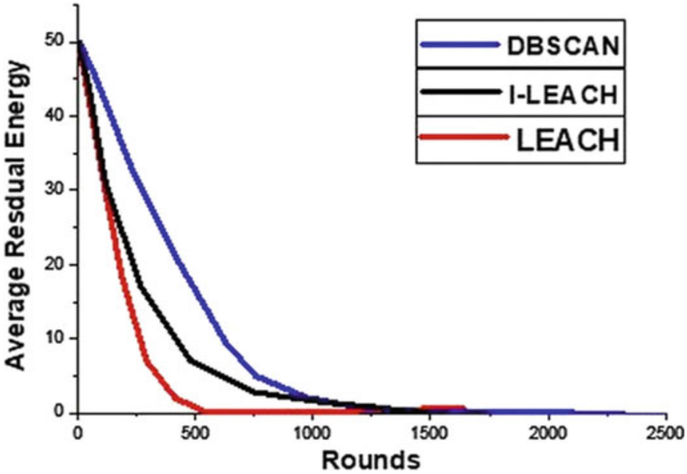


Fig. 3 Network performance **a** average residual energy

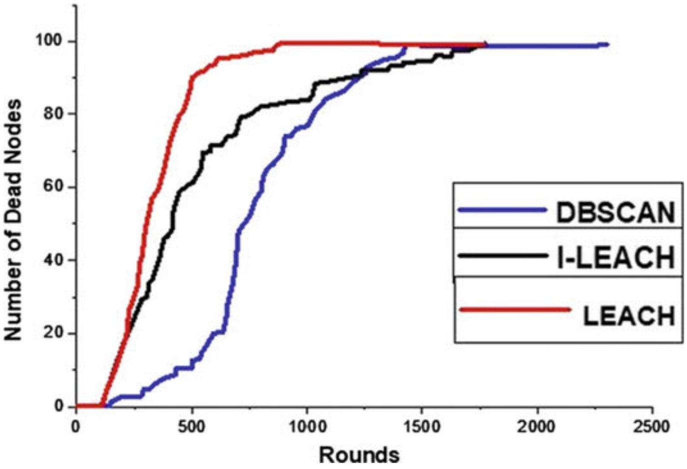


Fig. 4 Lifetime metrics: Dead nodes

whole situation for the maximization of the network lifespan, merely due to assigning various energy levels for differing communicating modes in the network.

The behavior of the algorithm is different in comparison of lifespan metrics with regards to the "first node dead (FND)" and "last node dead (LND)," as illustrated in Table 2. Simulation is performed with regards to three different areas (100, 200, and 400) m^2 with a network that is poorly to richly deployed with sensors. Keeping the initiated energy at (0.5 J) on a poorly deployed area of 100 m^2 improves the network timespan to 1.16, 1.36, and 1.3 for I-LEACH, LEACH, and CPCHSA, respectively. There is a positive relationship between the lifespan and number of nodes (keeping the area and energy constant), as it increases to 1.04, 1.87, and 1.19 times the value for I-LEACH, EECS, and LEACH, respectively. Doubling the area to 100 nodes results in a lifespan of 1.16, 1.3, and 1.47 times, the value for I-LEACH, ModLeach, and LEACH, respectively. At 400 nodes, the lifespan is 1.23, 1.4, and 1.61 times the value for I-LEACH, ModLeach, and LEACH subsequently. With the initiated energy of 1 J for 1000 nodes, the lifespan is 1.4, 1.4, and 3.5 times the value for I-LEACH, LEACH, and EECS. Comparing the DBSCAN to I-LEACH and LEACH over a wider area of 400 m^2 shows the increase of the stability period by (1.23, 2.6) and (1.28, 1.7), and (1.58, 1.7) times for 100, 400, and 1000 nodes, respectively. One can therefore draw the conclusion that the suggested protocol proves a more favorable performance with both smaller and larger areas, regardless of whether the networks were defectively or completely covered with nodes.

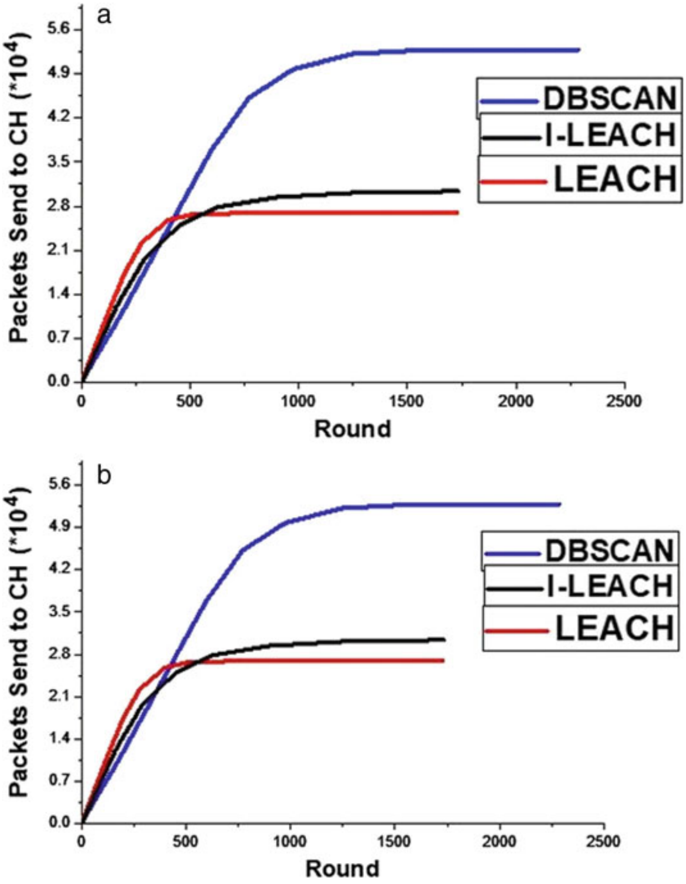


Fig. 5 Sent packets **a** to BS, **b** to CH

6 Conclusion

The clustering algorithms still represent a significant part in the field of wireless sensor network, and it gets a great consideration by many researchers in the current world. This article suggested a distributed density-based spatial clustering of applications with noise protocol for extending the lifetime of wireless sensors of Internet of Things networks. This protocol is distributed on every Internet of Things sensor device, and the sensor devices are cooperated based on the density-based spatial clustering of applications with noise algorithm to form the clusters in the network in a distributed way. The efficient periodic cluster head strategy is introduced based on certain criteria like remaining power, a number of neighbors, and the distance for each node in the cluster. The cluster head will be elected periodically and in a distributed way so as to the consumed power is balanced in the sensors inside each cluster. The

Table 2 Different scenarios for the network lifetime

Nodes	Energy	Area	LEACH		ILEACH		DBSCAN		Other protocols	
			FDN	LDN	FDN	LDN	FD	LD	FDN	LDN
100	0.5	100	980	1450	1050	1700	1150	1972	600 [45]	1500 [45]
		200	780	1150	850	1450	676	1685	200 [46]	1300 [46]
		400	100	800	98	1700	194	2092	–	–
400	0.5	100	1000	1500	1100	1700	1171	1778	820 [47]	950 [47]
		200	850	1300	900	1700	681	2093	190 [46]	1500 [46]
		400	100	1200	100	1600	316	2049	–	–
1000	1	100	2000	2700	2000	2700	2005	4095	–	–
		200	1700	2600	1700	2650	1121	3651	810 [47]	1050 [47]
		400	300	2500	300	2700	842	4284	–	–

proposed protocol is evaluated and compared to two existing methods using several performance metrics like the resting energy, network lifetime, and cluster head count, etc. The comparison results show that the suggested protocol can preserve power and improve the network lifetime better than other approaches.

References

1. Alhussaini R, Idrees AK, Salman MA (2018) Data transmission protocol for reducing the energy consumption in wireless sensor networks. In: New trends in information and communications technology applications, pp 35–49 (2018)
2. Idrees AK, Deschinkel K, Salomon M, Couturier R (2017) Multiround distributed lifetime coverage optimization protocol in wireless sensor networks. J Supercomput 74(5):1949–1972
3. Raj JS (2019) QoS optimization of energy efficient routing in IoT wireless sensor networks. J ISMAC 1(01):12–23
4. Idrees AK, Harb H, Jaber A, Zahwe O, Taam MA (2017) Adaptive distributed energy-saving data gathering technique for wireless sensor networks. In: 2017 IEEE 13th ınternational conference on wireless and mobile computing, networking and communications (WiMob)
5. Duraipandian M, Vinothkanna R (2019) Cloud based Internet of Things for smart connected objects.J ISMAC 1(02):111–119
6. Harb H, Idrees AK, Jaber A, Makhoul A, Zahwe O, Taam MA (2018) Wireless sensor networks: a big data source in ınternet of things. Int J Sens Wireless Commun Control 7(2)
7. Chamam A, Pierre S (2016) A distributed energy-efficient clustering protocol for wireless sensor networks. Comput Electr Eng 36(2):303–312
8. Fotouhi H, Alves M, Zamalloa MZ, Koubaa A (2014) Reliable and fast hand-offs in low-power wireless networks. IEEE Trans Mob Comput 13(11):2620–2633
9. Ranjan NM, Prasad RS (2018) LFNN: Lion fuzzy neural network-based evolutionary model for text classification using context and sense based features. Appl Soft Comput 71:994–1008

10. Geeta D, Nalini N, Biradar RC (2013) Fault tolerance in wireless sensor network using hand-off and dynamic power adjustment approach. J Netw Comput Appl 36(4):1174–1185
11. Behera TM, Samal UC, Mohapatra SK (2018) Energy-efficient modified LEACH protocol for IoT application. IET Wireless Sens Syst 8(5):223–228
12. Heinzelman W, Chandrakasan A, Balakrishnan H (2002) An application-specific protocol architecture for wireless microsensor networks. IEEE Trans Wireless Commun 1(4):660–670
13. Singh SK, Kumar P, Singh JP (2017) A Survey on successors of LEACH protocol. IEEE Access 5:4298–4328
14. Mahapatra RP, Yadav RK (2015) Descendant of LEACH based routing protocols in wireless sensor networks. Procedia Comput Sci 57:1005–1014
15. Heinzelman WR, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd annual Hawaii International conference on system sciences, Maui, HI, USA, vol 2, p 10
16. Rao PCS, Jana PK, Banka H (2016) A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks. Wireless Netw 23(7):2005–2020
17. Shankar A, Jaisankar N, Khan MS, Patan R, Balamurugan B (2019) Hybrid model for security-aware cluster head selection in wireless sensor networks. IET Wireless Sensor Syst 9(2):68–76
18. Li J, Jiang X, Lu I-T (2014) Energy balance routing algorithm based on virtual MIMO scheme for wireless sensor networks. J Sens 2014:1–7
19. Jan MA, Nanda P, Usman M, He X (2017) PAWN: a payload-based mutual authentication scheme for wireless sensor networks. Concurr Comput Pract Exp 29(17)
20. Purohit R, Bhargava D (2017) An illustration to secured way of data mining using privacy preserving data mining. J Stat Manag Syst 20(4):637–645
21. Ester M, Kriegel HP, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of 1996 ınternational conference on knowledg discovery and data mining (KDD ’96), pp 226–231
22. Han D, Agrawal A, Liao W-K, Choudhary A (2016) A novel scalable DBSCAN algorithm with spark. In: IEEE ınternational parallel and distributed processing symposium workshops (IPDPSW)
23. Idrees AK, Al-Yaseen WL, Taam MA, Zahwe O (2018) Distributed data aggregation based modified k-means technique for energy conservation in periodic wireless sensor networks. In: 2018 IEEE middle east and north africa communications conference (MENACOMM)

Hardware Implementation of Automatic Power Source Controller Cum Power Optimizer

A. Kunaraj, J. Joy Mathavan, and K. G. D. R. Jayasekara

Abstract A device that can switch from one source of power to the other or add on two power sources based on the real-time energy requirement. Power can be generated by various sources. Solar, wind, diesel generator, and electric board main grid supply are the four sources of energy tested here. The number of sources can be more based on user requirements. If the demand exceeds the supply, the device is programmed through AT Mega 2560 microcontroller to add more than one power source to meet the excess demand. With the help of electrical parameters, the power optimizer automatically switches and controls the power source based on the power consumption. If the demand exceeds the total supply, the device is programmed to stop certain devices until the consumption becomes normal.

Keywords Power optimizer · Power source controller · Solar power · CEB grid

1 Introduction

In the developed world, people use innovative ideas to invent new electrical/electronic appliances to do work more efficiently and effectively. The demand for electricity keeps on increasing when everyone starts using electrical appliances. The objective of this research work is to propose an automatic power controller cum power optimizer to switch the power from one source to another or to add up more than one power source based on demand for power. A power generation system should be able to produce enough power using various sources. This device is designed to go for the energy produced by cheap and available sources first, and if the demand exceeds the supply, the device will switch to the next available power source. It can be

---

A. Kunaraj · J. Joy Mathavan (✉) · K. G. D. R. Jayasekara
Faculty of Technology, University of Jaffna, Jaffna, Sri Lanka
e-mail: joymathavan1991@gmail.com

A. Kunaraj
e-mail: kunaraj12@gmail.com

K. G. D. R. Jayasekara
e-mail: darshanaruma43@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_3

used in situations where the electrical energy consumption is high and fluctuating in an unpredictable manner. Mainly, four power systems can be connected with this proposed device. The power optimizer can measure the power requirement of the consumer and take a suitable real-time decision automatically to process and switch to the next available power source. In most of countries, the energy produced by private companies or by individuals, if it is excess, can be supplied to the main grid line of government, based on the energy requirement. The sources of power generation may be costly or cheap depending on the availability of those sources. If a cheap and available energy resource are used to produce maximum power, a large amount of money could be saved. Solar and wind power are cheap renewable energy sources. In contrast, coal and diesel are non-renewable sources and they harm the environment as well.

Based on the power requirement of the industries and domestic users, there is no proper power optimizing devices for power management. People usually use the main grid line to their electrical energy requirements. When there is a power cut or low voltage supply, they use solar energy, if they have it. If there are only two power sources, it can be switched at least manually. But when the number of power sources increases, it is very difficult to manage the switching between power sources. This project considers a new device to control four power sources manually and automatically. The automatic mode operated based on the power requirement of the industries as well as domestic purposes. In case if the power consumption is high due to the heavy use of electrical appliances, it can be controlled using the wireless control method. The control method can either be switching to the next available power source or turning off certain devices for a limited period.

Solar energy, which is obtained from the sun, can be converted directly to solar energy using photovoltaic solar cells [1]. Ricardo Orduz et al. discussed the development of the PV cell and maximum power point tracking (MPPT) system. Generally, the solar panel combination has one MPPT; but in this work, each PV panel has an individual MPPT module and all the MPPT modules connected [2].

Wind energy is also a readily available and eco-friendly source of energy [3]. The usage of wind energy to produce electrical energy is first implemented in America in 1887 [4]. Various developments in the selection of composite materials for wind turbine blades are still in progress [5]. Sara Mac Alpine et al. discussed mitigating the losses related to non-uniform operating conditions in grid-tied photovoltaic arrays [6]. Elkamouny et al. mentioned in their research work about the recent developments in the technologies associated with capturing solar energy through solar cells [7]. Sanz et al. researched DC-DC converter PV system architecture and they have mentioned that almost all the mismatching losses between modules are eliminated in their outcome, and the energy output has increased [8]. Muthamizhan mentioned that coupled inductor and switched capacitor technologies were used to get high voltage gain when considering DC distribution system with solar power optimizer [9]. Salpe performed research in solar power optimizer to get maximum energy from a photovoltaic panel and send this energy to a DC microgrid. A coupled inductor and switched capacitor are used to increase the voltage gain [10]. Sivakumar explained about boost converter using SIC diodes for PV applications. The converter designed here has two switching cells to distribute the output current [11]. Haoxiang et al. focused on developing a multi-objective optimization algorithm (MO-OPA) for power management in the radio networks. This method helps to reduce power consumption by minimizing the delay in communication [12]. Smys developed ad hoc networks in various traffic conditions [13]. Most of the researches mainly focus either on one or two power sources. One among solar panels or wind or hybrid sources and the other one is the AC grid line. These researchers tried to mitigate the existing issues in the available two power source switching systems. In this current research work, the necessary power is supplied by a design of four power sources that can switch to manage the power requirement.

2 Methodology

The automatic power optimizer cum power source controller can be divided into two parts namely the master module and the slave module. All the slave modules can be controlled by the signals of the master module. The number of slave modules can be increased according to the user requirement. This research work aimed to switch four power sources based on the energy requirement. The master controller mainly focuses on increasing the use of renewable energy sources and if the demand exceeds only, it will switch to the other source. Solar energy is an environmentally comfortable, renewable, limitless [14], and cost-effective [15] source among all other alternative sources used here. Solar energy will become an economical source of energy in the coming years and developing good technology for solar cells, reduction of cost, and efficiency in an application [16, 17]. So, the use of solar power is prioritized here according to the control given to the master controller. Usually, if the solar power supply is not enough to operate appliances which need high power, like refrigerator, cooler or motor, those appliances should be turned off.

But in this device, the power derived from the next available source will be added up with the existing power source and the continuous working of those appliances is ensured. High power electrical equipment is controlled by a slave module which automatically switches off certain devices, in case, if the total power supplied by all the available power sources cannot meet the demand. The slave modules are connected with the master module by the radio signal. If the demand exceeds the supply, the slave module sends an RF signal to the master module, and the master module gives the command to switch up or shut down that particular device. Within 100 m^2, the master module and other slave modules can communicate with each other. The workflow diagram of the system is shown in Fig. 1, and the block diagram of the system is shown in Fig. 2.

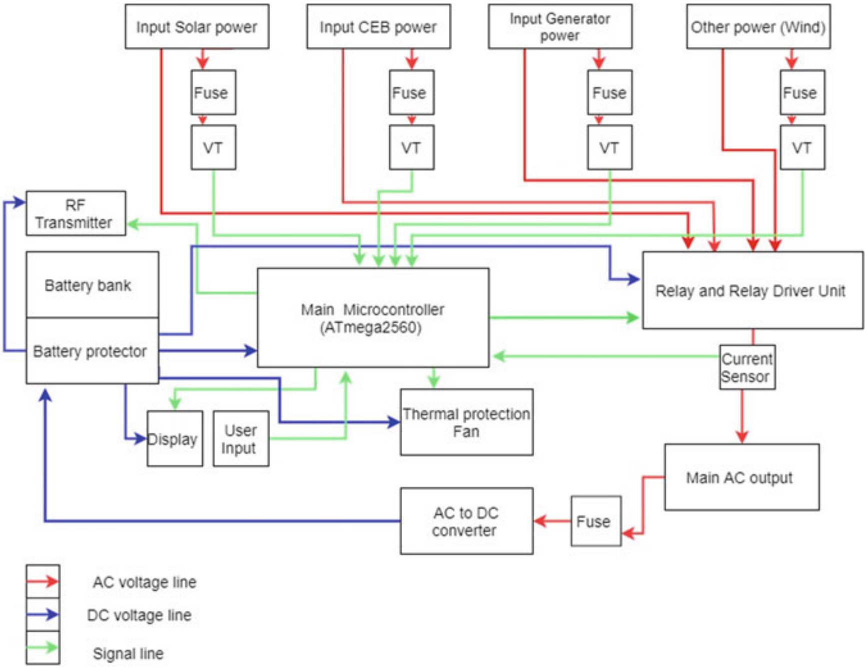


Fig. 1 Conceptual overall workflow diagram

2.1 AT Mega 2560 Microcontroller

AT Mega 2560 microcontroller is selected in this project since a large number of input and output are expected to be controlled by the microcontroller. AT Mega 2560 microcontroller is the big member of AT Mega series, and it has more number of I/O pins. The specifications of it is shown in Table 1. There are 16 analog pin in AT Mega 2560 microcontroller for analog processing, 54 pins for the general digital input–output processing, and 15 pins for the Pulse Width Modulation (PWM) processing.

In this research, work pins of all those categories are used to get inputs and give outputs. The keypad and wireless communication are connected as inputs such as the relay driver circuit, LED indicator, and LCD are connected as output in digital input–output pins. The current sensor is connected as input, and the VT transformer input is connected as output with Pulse Width Modulation (PWM) pin. Generator and solar data input are connected as the input to the system protection (Fan) are connected as output with analog pins.

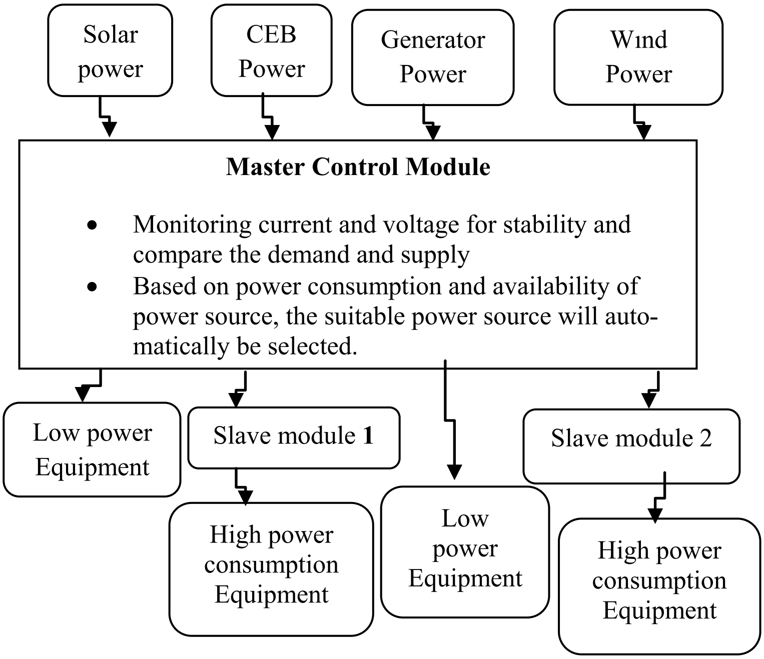

Solar power
CEB Power
Generator Power
Wınd Power
Master Control Module
Monitoring current and voltage for stability and compare the demand and supply
Based on power consumption and availability of power source, the suitable power source will automatically be selected.
Low power Equipment
Slave module 1
Low power Equipment
Slave module 2
High power consumption Equipment
High power consumption Equipment

Fig. 2 Power optimizer full system block diagram

Table 1 Specifications of AT Mega 2560 microcontroller

Parameters	Range
Microcontroller	At mega 2560
Operating voltage	5 V
Digital I/O pins	54
Analog pins	16
PWM pins	15
Crystal oscillator	16 MHz
Current rating per I/O pin	20 mA

2.2 Relay Controller Module

The relay module is an important component in this equipment since all the processing is done on the AC power sources through the relay module. The relays are driven by relay driver circuit, and relay driver circuit is controlled by the main microcontroller (Fig. 3).

Based on the power requirement, the relay controller module switches from one source of power to the other or add up with the other and it is shown in Fig. 3. Four power sources namely solar, Ceylon electric board (CEB), generator, and wind are

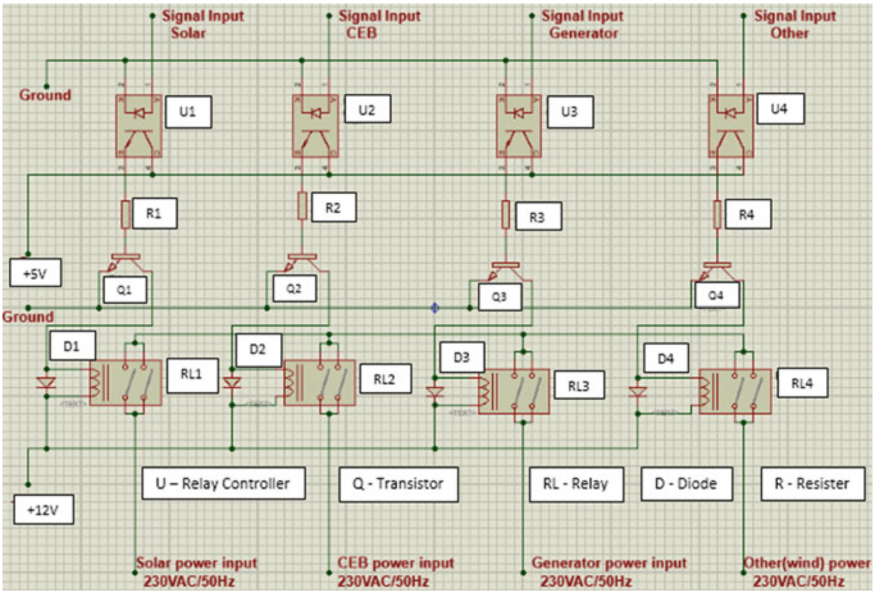


Fig. 3 Relay driver circuit diagram

occupied. The last source is mentioned as “other” in the diagram since any other source of power can be annexed with the existing design. Depend on the power demand, the microcontroller sends the signal to the relay circuit and energizes the relevant power source. When the relay is operated, back EMF would be generated and this signal affects the performance of the microcontroller since the relay is connected directly with the microcontroller. The rectifier diode is connected with the relay driver to send back EMF to the ground and prevent the intervention of back EMF in the performance of the microcontroller.

2.3 AC to DC Converter

Since all the processing inside the master controller module is occurring in DC voltage, the AC voltage needs to be converted to DC voltage by the AC to DC converter as shown in Fig. 4. Firstly, 230 VAC is reduced to 12 VAC, and after rectification and smoothening, 12 VDC output is obtained. 7805 IC is used to regulate the voltage from 12 to 5 VDC since 5 VDC is required to operate the microcontroller. One microfarad capacitor is used to get smoothened 5 VDC output.

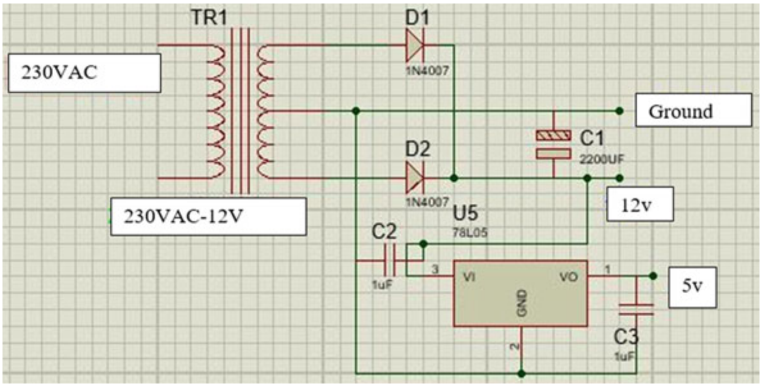


Fig. 4 Circuit diagram of AC to DC converter

2.4 Battery Bank

A battery bank is used as a reserve power source for the internal operation of the master controlling module. If the main power source failed to provide power to the device, the battery bank will provide it. If there is any interruption in the supply power to the master module, the whole system will be shutdown. Since the proper functioning of the device is needed always to switch the power, the battery bank is used as a backup and it is used to store the electricity to provide to the master controlling module. Three Lithium-ion batteries each of 3.7 V and 4000 mAh as shown in Fig. 5 are connected in series to give an approximate of 12 V output.

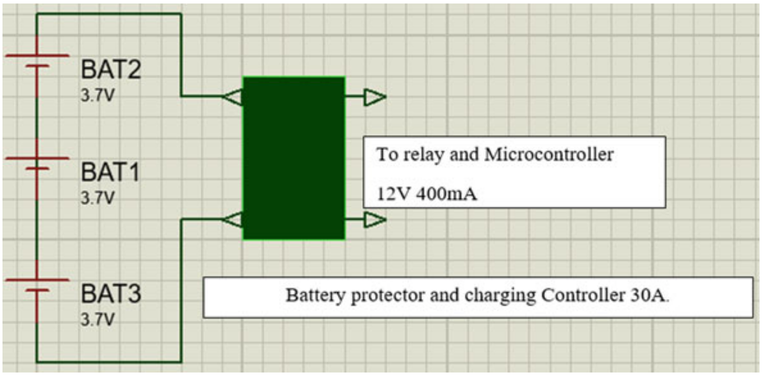

BAT2
3.7V
BAT1
3.7V
BAT3
3.7V
To relay and Microcontroller
12V 400mA
Battery protector and charging Controller 30A.

Fig. 5 Circuit diagram of battery bank

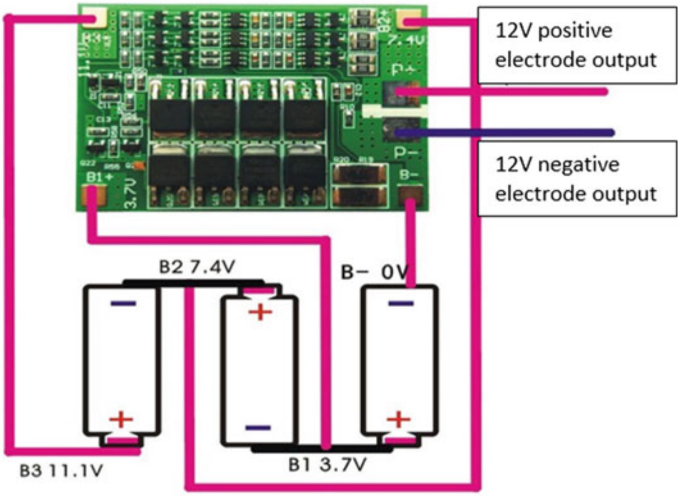


Fig. 6 Charging controller

2.5 Charging Controller and Battery Protector

The charging controller is designed in a way as shown in Fig. 6. Usually, the batteries are charged in a way that all batteries are connected in series. In case if one battery is damaged, none of the other batteries would be charged. Therefore, to prevent such happening and to protect the battery and other equipment connected with the battery, the individual charging of each battery is designed. In this way, the battery is charged individually and the lifetime of the battery will also be high.

2.6 Current Sensor

The current sensor is used to detect the electric current in a circuit. This sensor generates a signal proportional to the current. The generated signal may be voltage or current or digital output. ASC712 IC is used as a current sensor.

It has 8 pins and the supply voltage to it is 5 V. One pinout is provided for the analog output. When current flow through the ASC712 IC, the IC output analog voltage change between 0 and 5 V. The variation is mentioned in Fig. 7.

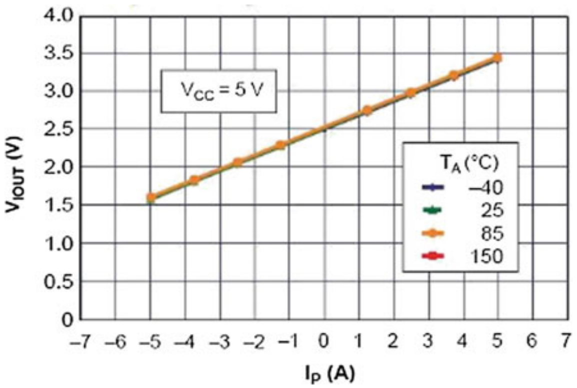


Fig. 7 Variation of current with voltage supply

2.7 Display Module

The display module shown in Fig. 8 is used to display the data needed by the user. 128 * 64 display is used in this project, and it has 20 pins which are also suitable for displaying the graphics. There are many Arduino libraries that can support this display module.

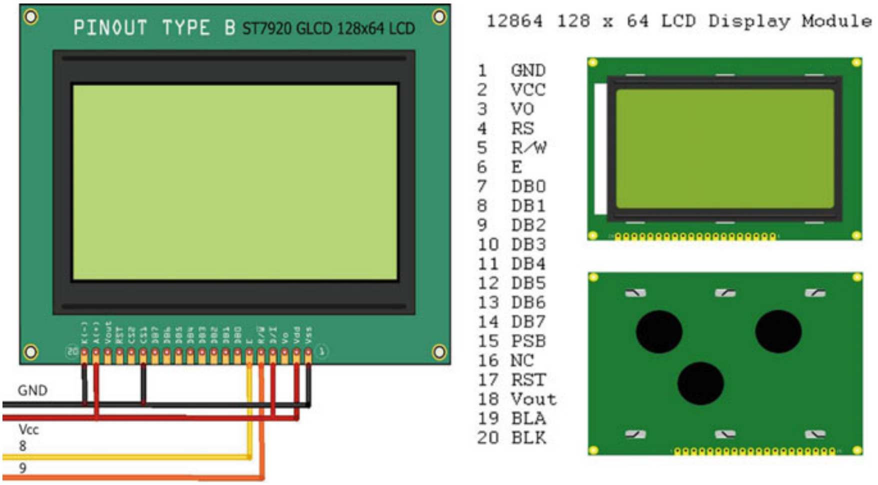
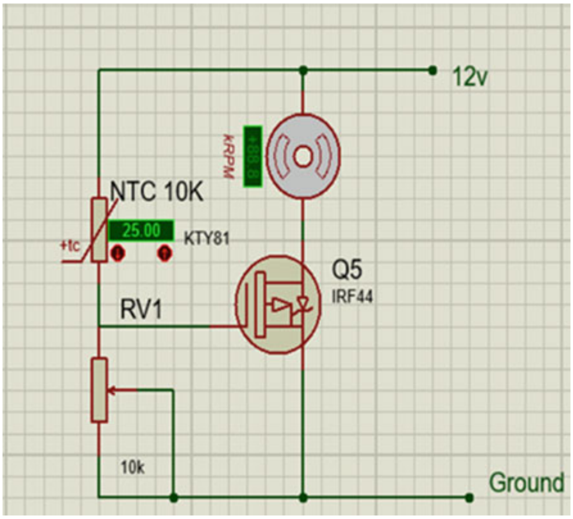


Fig. 8 Display module

Fig. 9 Cooling system



2.8 Thermal Protection to the System

The master controlling module is a multi-processing unit and VT transformers are installed inside the module. Usually, transformers heat up while operating. In order to prevent the heating up of the transformers, normally an in-built cooling system would be provided in most of the transformers. But VT transformers usually does not have a cooling arrangement. Therefore, a separate cooling system is provided in this research work as shown in Fig. 9. The temperature should be maintained at the proper level by the automatic cooling fan for the efficient functioning of the transformer. When the temperature increases beyond a certain point, the driver circuit of the fan identifies it based on thermistor reading because the resistance of the thermistor changes with temperature. NTC thermistor is used in this project. If the temperature increases, the resistance across the terminals would be reduced. In this condition, IRF44 MOSFET would be biased and current flows through the source to drain. This will switch ON the fan and thereby cool the system. Once the temperature reduces below the prescribed value, the fan automatically switches off.

2.9 Voltage Transformer

The voltage transformer is used to measure the AC voltage. The 230 VAC is converted to 6 VAC, and it is rectified and smoothened by 10 microfarad capacitors to get the regulated voltage of 5.1 V and this is the input voltage signal to the microcontroller. The primary coil side of the transformer is connected with 800 mA fuse in series

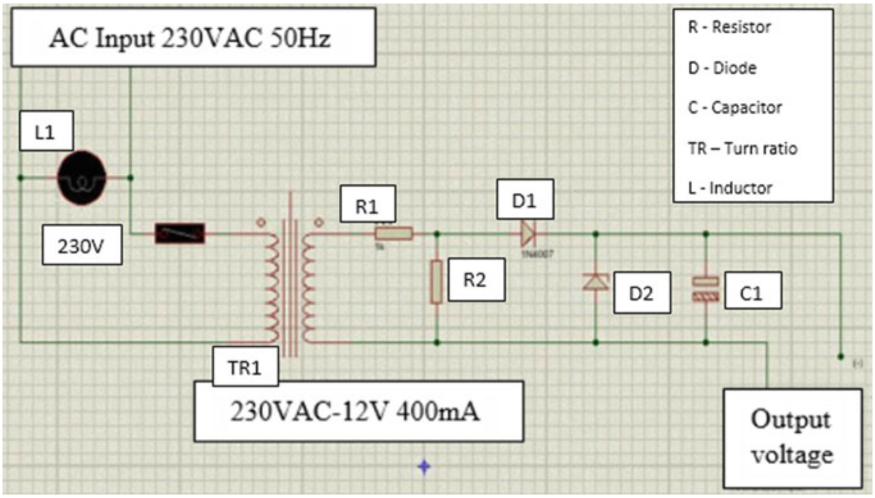


Fig. 10 Circuit diagram of voltage transformer

to protect the VT transformer from overvoltage. The circuit diagram of the VT transformer is shown in Fig. 10.

2.10 Data Access Port

When operating various energy sources, the master controlling module needs the amount of energy that can be supplied by each source. The data access ports are introduced to access the data of each power source. The solar power source produces solar radiation and heat. And, if it is an electricity generator calculate the amount of fuel left and if it is a wind power source determine the speed of wind, etc. Not only the basic data, but also the data like inverter mode and PV voltage for solar source, the pulse of starting motor and generator temperature for generator source, and all the data necessary for the user would be provided by data access port. These data are provided to the master module through a data access port. 5.1 V zener diode is used in this module to limit the input signal up to 5 V since the input data to solar and generator sources should be around 5 V since the operating voltage of the microcontroller is 5 V. If more than 5 V flows through the microcontroller, it will be damaged. The circuit diagram and schematic diagram of the data access port are shown in Figs. 11 and in 12, respectively.

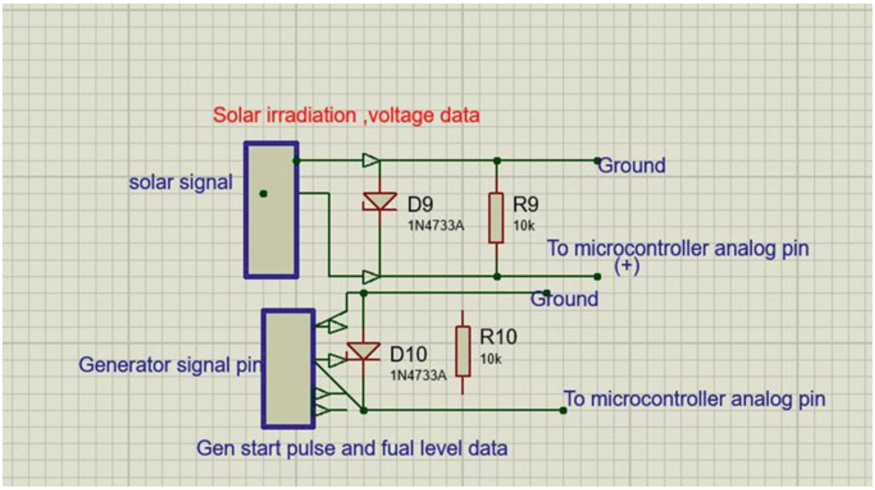

Solar irradiation ,voltage data
solar signal
Ground
D9
1N4733A
R9
10k
To microcontroller analog pin
(+)
Ground
R10
10k
Generator signal pin
D10
1N4733A
To microcontroller analog pin
Gen start pulse and fual level data

Fig. 11 Circuit diagram of data access port

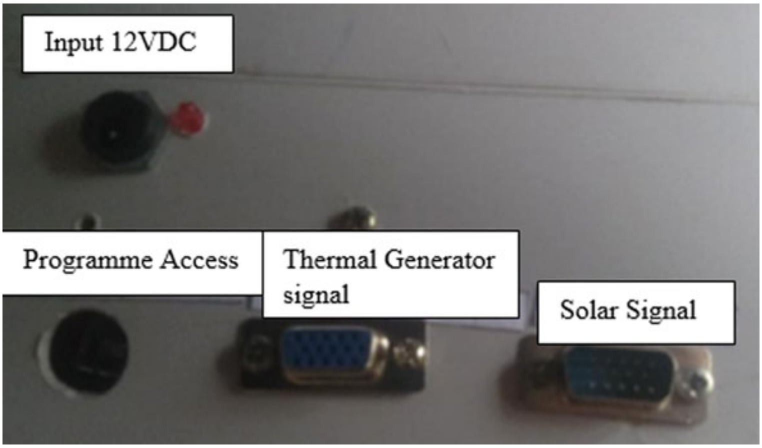


Fig. 12 Diagram of data access ports in the device

2.11 Fuse

Each circuit has a series of fuses connected with it to protect the circuit from the threat of overvoltage or short circuit as shown in Fig. 13. The damage occurs in the circuit of one source, damaging the other circuit is also prevented by the introduction of fuses. The fuses can easily be replaced since the fuse holders are mounted outside the module case. Connectors are used to connect each input and output wires, and it can easily be removed and safely be connected with the master module. Each connection have a separate terminal as shown in Fig. 14.

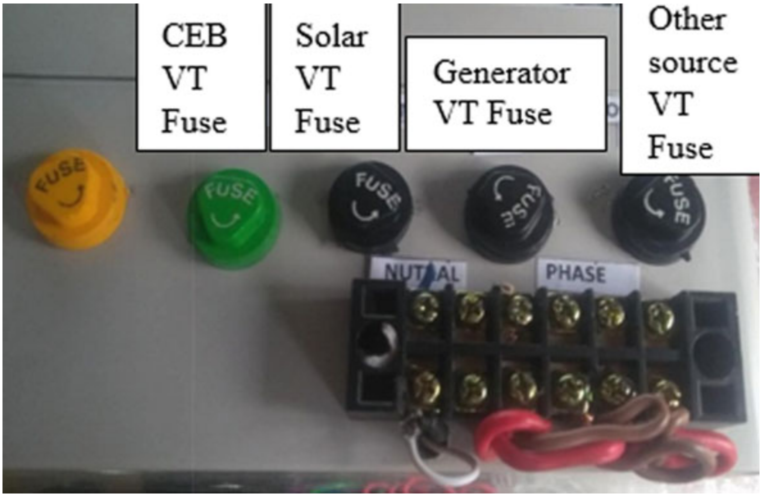


Fig. 13 Arrangement of fuses

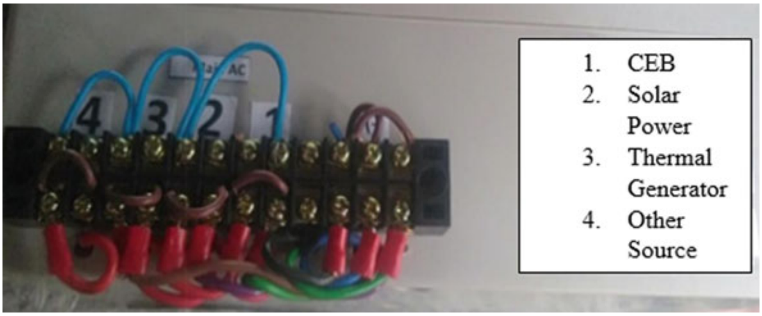


Fig. 14 Diagram of connector terminal bar

2.12 *Slave Module*

The slave module could be located in different locations that are connected with the master module. Heavy equipment that consumes high power is connected with the main output AC line through a slave module as shown in Fig. 15. In case if the combined power supply by all four sources cannot meet the demand, the slave module needs to turn off some equipment connected through it with the main AC line. The power is supplied to the slave module as shown in Fig. 16.

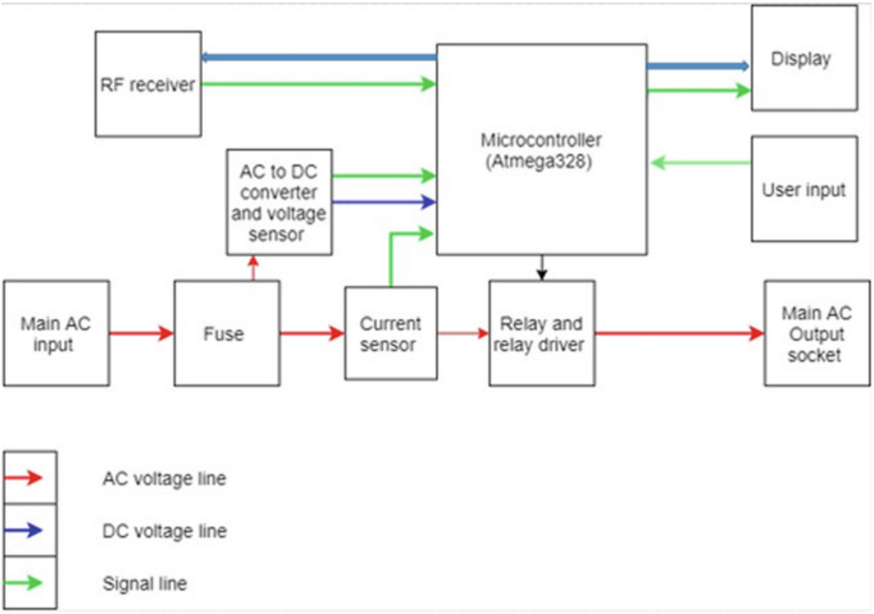


Fig. 15 Block diagram of slave module

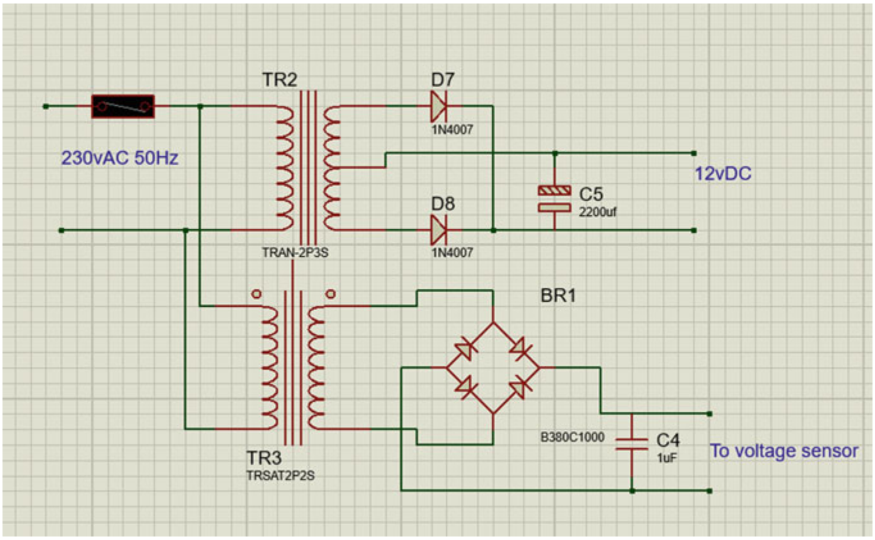


Fig. 16 Power supply for slave module

3 Results and Discussion

Wireless connection is established between the master controlling module and the slave module. A cheap and reliable method of wireless connection are a radio signal. 433 MHz transmission module is used since it can cover a large working area of 100 m and it is comparatively cheap. The transmitter module is in the master controlling module and the receiver module is in each slave module. When four power sources are connected with the master controller module, it checks the availability of the power sources, reads the voltages and currents of all power sources, and also reads the power consumption of the main AC output. The master controller module always tries to connect with the solar power sources at the first attempt.

At first, the master control module checks the availability of solar power. If solar power is available, the solar power source would be selected and connected with the main AC output. If the solar power source is not available, on the next turn, the master control module checks the availability of CEB power. If CEB power is available, the main AC output would be connected with the CEB power source. If the CEB power source is not available, the master controller module on its third run checks the generator fuel level. If the fuel level is sufficient, the master control module sends the signal to the generator to start. So, the generator would be started and generator power can be connected with the main AC output. During peak hours, the solar power is not sufficient most of the time. Therefore, the master controller module always checks the availability of CEB and generator.

In case if the combined supply of all four power sources together is not enough during peak hours of power consumption, the master controller module sends the signal to the slave module to cut-off the power from equipment like air conditioner which consumes high power. Equipment consume less power to keep on working while equipment consumes high power would be shut down until the balance between supply and demand arise. The basic idea of this device is to increase the usage of easily available renewable power sources instead of conventional government power supply. The prevention of environmental pollution and the economic growth of the country is also expected to be addressed through this research work. For example, coal and diesel power which are used in thermoelectric plants are polluting the environment and, it has the threat of extinguishing soon. In most of the countries, coal and petroleum products are imports. Renewable resources like solar, tidal, wind, and hydro energy are environment-friendly and readily available energy sources (Fig. 17).

4 Conclusion and Future Scope

There are two modules developed in this design, one is the master control module and the other one is the slave module. The master control module has control over the power sources and the parameters of all individual power sources. The slave module is used in a way that the heavy equipment which consumes high power is

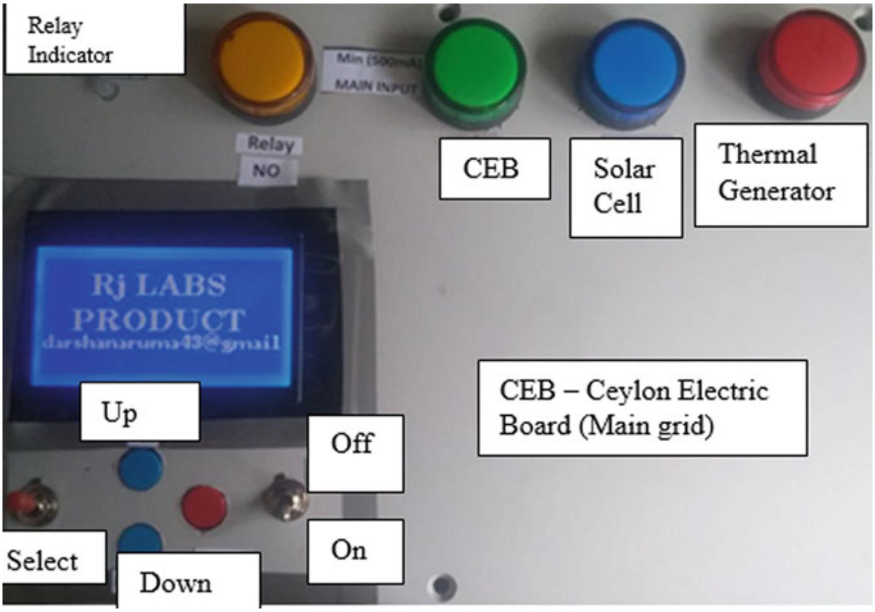


Fig. 17 Real-time working model of the proposed system

connected through the slave module with the main AC grid line. If power production from one source is not enough, the master module programmed to add up more than one power source to provide the necessary power. If the combined supply of all four power sources becomes lower than demand and the power supply from the source is detected by the master controlling module as low, it sends a signal to the salve module to switch off certain devices based on the output AC voltage. Similarly, if the master controlling module detects sufficient power from the sources, then it sends the signal to all the slave modules to switch ON the output AC supplies which are closed.

The working of the power optimizer cum automated power sources controller is based on the availability of the power sources and power demand on the main AC output. As a development of this controller, the addition of the power consumption for every month is expected to be calculated daily based on KW/h to find a graph of peak hours. If it is found, then the master controller can select the power sources according to this graph as a predetermined function. And also it would be very useful in the industrial point of view since the industries run cyclic workloads on their day-to-day functioning. If the heavy-duty hours of the industries are found accurately, this device can be programmed appropriately and maintenance of the device and the system will also be easy.

References

1. Shaikh MRS, Waghmare SB, Labade SS, Fuke PV, Tekale A (2017) A review paper on electricity generation from solar energy. Int J Res Appl Sci Eng Technol 5(IX):1884–1889
2. Orduz R, Solórzano J, Egido MÁ, Román E (2011) Analytical study and evaluation results of power optimizers for distributed power conditioning in photovoltaic arrays. Fundación Tecnalia Research and Innovation, Energy Unit, Edit. 700 Derio 48160, Spain
3. Suresh Babu K, Raju S, Srinivasa Reddy M, Nageswara Rao DN (1887) The material selection for typical wind turbine blades using MADM approach & analysis of blades. In: MCDM 2006, Chania, Greece, June 19–23
4. Kumar A, Dwivedi A, Paliwal V, Patil PP (2014) Free vibration analysis of Al 2024 wind turbine blade designed for Uttarakhand region based on FEA. Procedia Technol 14:336–347
5. Mathavan JJ, Patnaik A (2020) Development and characterization of polyamide fiber composite filled with fly ash for wind turbine blade. In: Emerging trends in mechanical engineering. Springer, Singapore, pp 131–139
6. MacAlpine SM, Erickson RW, Brandemuehl MJ (2013) Characterization of power optimizer potential to increase energy capture in photovoltaic systems operating under nonuniform conditions. IEEE Trans Power Electron 28(6):2936–2945. https://doi.org/10.1109/TPEL.2012.2226476
7. Elkamouny K, Lakssir B, Hamedoun M, Benyoussef A , Mahmoudi H (2017) Simulation, design and test of an efficient power optimizer using DC-DC interleaved isolated boost PV-micro inverter application. In: 2017 14th ınternational multi-conference on systems, signals & devices (SSD), Marrakech, pp 518–525. https://doi.org/10.1109/SSD.2017.8167019
8. Sanz A, Vidaurrazaga I, Pereda A, Alonso R, Román E, Martinez V (2011) Centralized vs distributed (power optimizer) PV system architecture field test results under mismatched operating conditions. In: 2011 37th IEEE photovoltaic specialists conference, Seattle, WA, pp 002435–002440. https://doi.org/10.1109/PVSC.2011.6186440
9. Muthamizhan T (2016) Performance analysis of solar power optimizer for DC distribution system. Int J Adv Res Electr Electron Instrum Eng 5(8):6708–6715
10. Salpe AD (2016) Design, analysis, of efficient solar power optımizer for DC mıcro grid system. Int J Adv Res Sci Eng 5(6):182–189
11. Sivakumar K (2014) Implementatıon of interleaved boost converter using SIC diodes in residential PV pre-regulator application. Int J Adv Eng Technol 6(6):2537–2547
12. Haoxiang W (2019) Multi-objective optimization algorithm for power management in cognitive radio networks. J Ubiquit Comput Commun Technol (UCCT) 1(02):97–109
13. Smys S, Josemin Bala G, Jennifer S (2010) Mobility management in wireless networks using power aware routing. In: 2010 International conference on intelligent and advanced systems, pp. 1–5. IEEE
14. Chu Y, MeisenP (2011) Review and comparison of different solar energy technologies. Report of Global Energy Network Institute (GENI), Diego
15. Choubey PC, Oudhia A, Dewangan R (2012) A Review: solar cell current scenario and future trends. Recent Res Sci Technol 4:99–101
16. Wall A (2014) Advantages and disadvantagesofsolarenergy. Process Industry Forum, 7 Aug 2013. Web, 2 Feb 2014
17. Bagher AM, Vahid MMA, Mohsen M (2015) Types of solar cells and application. Am J Opt Photon 3:94–113

Texture-Based Face Recognition Using Grasshopper Optimization Algorithm and Deep Convolutional Neural Network

Sachinkumar Veerashetty and Nagaraj B. Patil

Abstract Face recognition is an active research area in biometric authentication, which has gained more attention among researchers due to the availability of feasible technologies, including mobile solutions. However, the human facial images are high dimensional, so the dimensionality reduction methods are often adapted for face recognition. However, the facial images are corrupted by the noise and hard to label in the data collection phase. In this study, a new GOA-DCNN model is proposed for face recognition to address those issues. Initially, the face images are collected from two online datasets FEI face and ORL. Next, modified local binary pattern (MLBP) and speeded up robust features (SURF) are used to extract the feature vectors from the collected facial images. The extracted feature values are optimized using grasshopper optimization algorithm (GOA) to decrease the dimensionality of data or to select the optimal feature vectors. At last, deep convolutional neural network (DCNN) was applied to classify the person's facial image. The experimental result proves that the proposed model improved recognition accuracy up to 1.78–8.90% compared to the earlier research works such as improved kernel linear discriminant analysis and probabilistic neural networks (IKLDA + PNN) and convolutional neural network (CNN) with pre-trained VGG-Face.

Keywords Deep convolutional neural network · Face recognition · Grasshopper optimization algorithm · Modified local binary pattern · Speeded up robust features

1 Introduction

In recent years, face recognition plays a vital role in a biometric authentication system that is applied in many applications such as law enforcement, access control, video

---

S. Veerashetty (✉)
Appa Institute of Engineering and Technology, Kalaburagi, India
e-mail: sveerashetty@gmail.com

N. B. Patil
Government Engineering College, Yaramaras, Raichur, India
e-mail: nagarajbpatil1974@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_4

surveillance, and access control [1]. Due to the rapid growth of electronic equipment techniques, a large number of face images are captured using cell phones and cameras. Hence, the image-based facial recognition becomes essential in real-world applications. The human facial images have high dimensionality that leads to the curse of dimensionality and computational complexity in the face recognition system [2]. Meanwhile, the accuracy of the face recognition system is compromised dramatically in real-world applications by inter and intraclass variations, due to facial expression, occlusion, aging effect, head pose variation, poor illumination, and low and blur resolutions [3—6]. To address these concerns, many dimensionality reduction methods have been developed based on unsupervised, supervised, and semi-supervised conditions. The unsupervised methodologies include principal component analysis [7], sparsity preserving projection, etc. The supervised methods include linear discriminant analysis [8], maximum margin criterion, etc. Still, the existing face recognition techniques are not fulfilling the requirements of real-world applications in the case of large scale data scenes. So, a superior optimization technique with a deep learning classifier (GOA-DCNN) is proposed in this study.

At first, the facial images are collected from two databases FEI face and ORL. Then, image quantization is carried out to improve the quality of the facial images. The image quantization technique reduces the number of colors utilized in the facial images, which is essential to display the images on devices like mobile phones, biometric attendance devices, etc. Besides, MLBP and SURF features are used to extract the feature vectors from the denoised facial images. In LBP, the sign vector is failed to extract the important texture feature vectors, so the rotation invariance, scale, and illumination vectors are calculated in MLBP for extracting all the texture feature vectors. The extracted features are high dimensions in nature, which is optimized by GOA to reduce the curse of dimensionality issues. The optimized features are classified by the DCNN classifier to classify the person's face images.

The paper is arranged in the following manner, a few recent research papers on face recognition is explained in Sect. 2. The proposed model with mathematical expressions is discussed in Sect. 3. Section 4 presents the quantitative and comparative results of proposed and existing models. Finally, the conclusion is described in Sect. 5.

2 Literature Survey

Ouyang et al. [9] presented a hybrid model improved kernel linear discriminant analysis (IKLDA) and probabilistic neural networks (PNNs) for facial recognition. Initially, the IKLDA method was adapted to reduce the dimension of the extracted features to retain the most relevant information about the facial images. Next, the PNN method was applied to solve the difficulties of facial recognition. The developed model (IKLDA and PNN) not only enhanced the recognition accuracy but also improved the overall computing efficacy. In this study, the performance of the developed model was validated on three databases like AR, YALE, and ORL. These databases comprise a wide range of face details, expressions, and scale degrees. The experimental results proved that the developed model achieved better recognition accuracy compared to the existing techniques. In facial image classification, PNN is slower than multi-layer perceptron networks, because it requires more memory space for storing the model.

Faraji and Qi [10] presented a multi-scale approach for facial recognition based on the maximum response filter. Initially, the facial images were scaled using a log function for compressing the brighter image pixels and expanded the darker image pixels. The multi-scale approach used a filter bank to reduce the illumination and enhanced the edges of the image. At last, an improved multi-scale gradient face method was used to capture different properties of the facial images to generate an illumination invariant feature representation. The developed multi-scale approach attained good performance related to other earlier methods of facial recognition. In this literature study, manual intervention is high during testing and training of the data that increase the computational time of the system.

Elmahmudi and Ugail [11] utilized a convolutional neural network (CNN) with a pre-trained VGG face model to extract the features and classify an individual's face. In this study, labeled faces in the wild (LFW) and Brazilian FEI datasets were used to evaluate the performance of the developed model. Simulation outcomes showed that the individual parts of the face such as nose, cheeks, and eyes were achieved a better rate of recognition. The conventional CNN does not encode the position and orientation of the face cues, and also, it is computationally expensive. Besides, Li and Suen [12] developed a new model for facial recognition by extracting the discriminate parts and dynamic subspace of the facial images. These parts represent the discriminative components and provide a recognized protocol to classify the facial images. In this study, the experiment was performed on three online datasets extended Yale B, ORL, and AR to validate the speed, robustness, and accuracy of the developed model. However, face occlusions, and variations are the major concerns in this study to develop a robust face recognition system.

Li et al. [13] implemented recurrent regression neural network (RRNN) for facial recognition. In the RRNN classifier, the encoder-decoder was a first unit, which was used to model sequential reconstruction. The second unit was utilized for constraining the global nature of the sequences, and the final unit was utilized to label the discriminative information. The experimental results proved that the RRNN classifier achieved better recognition results compared to the existing methodologies. However, there is a loss of information while embedding the low-resolution facial images in higher-level layer, where the RRNN classifier contains several a higher-level layers. Tang et al. [14] used LBP to extract the feature values from the collected face images. Then, ten CNN with five dissimilar network structures were employed for extracting the feature vectors for training that enhance the network parameter and classification result utilizing softmax function. At last, a parallel ensemble learning methodology was applied for generating the result of face recognition. As previously mentioned, CNN requires high graphics processing unit system for attaining better performance in face recognition, where it is highly expensive. To highlight these concerns, a new optimization technique with a deep learning classifier is proposed in this study for facial recognition.

3 Methods

The face recognition is an emerging research topic, which attracts more researchers in the field of pattern recognition and computer vision [15, 16]. In multimedia applications, face recognition has great potential, for instance, personal verification, video surveillance, digital entertainment, etc. Therefore, image-based face recognition is necessary for many real-time applications and it becomes a popular research topic in the area of facial recognition [17, 18]. In this research, the proposed GOA-DCNN model contains five steps: image collection from FEI face and ORL databases, image pre-processing (i.e., image quantization), feature extraction using MLBP, and SURF, feature optimization using GOA, and classification using DCNN. The workflow of the proposed model is presented in Fig. 1.

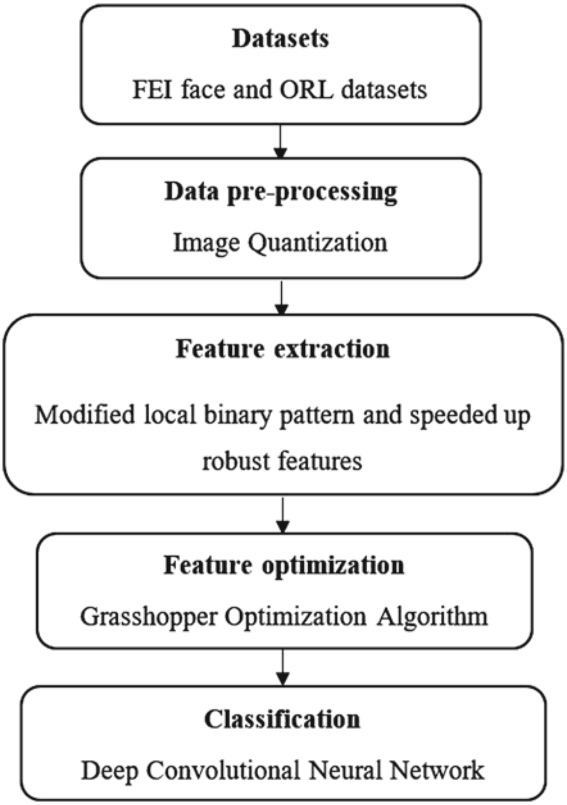

Datasets
FEI face and ORL datasets
Data pre-processing
Image Quantization
Feature extraction
Modified local binary pattern and speeded up robust features
Feature optimization
Grasshopper Optimization Algorithm
Classification
Deep Convolutional Neural Network

Fig. 1 Block diagram of proposed model

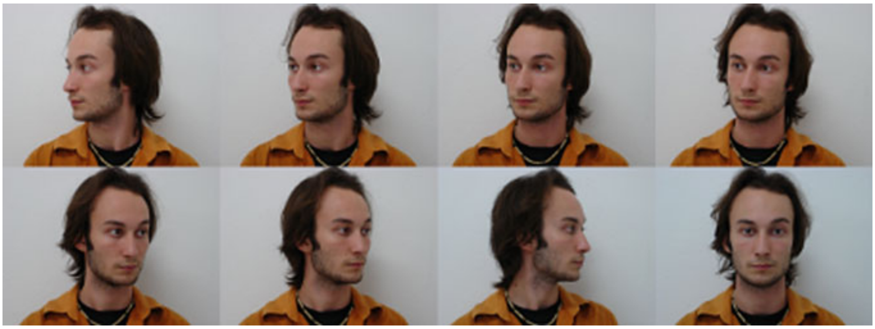


Fig. 2 Graphical illustration of FEI face dataset

3.1 Dataset Description

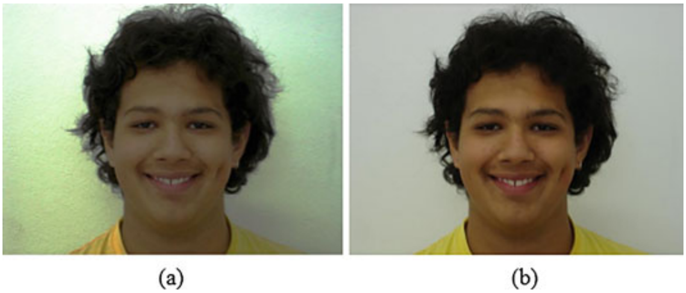
In this study, the input facial images are collected from two datasets FEI face and ORL. FEI dataset includes different facial images that are collected from the period of June 2005 to March 2006 at the artificial intelligence laboratory of the FEI in Brazil [19]. In the FEI dataset, the facial images are collected from 200 subjects; each subject includes 14 images and a total of 2800 facial images. In this dataset, all the facial images are colorful, those were taken under white homogenous background and the size of every facial image is 640 × 480 pixels. In this dataset, the facial images are collected from staff and students at FEI. Where the age ranges between 19 and 40 years old with a different hairstyle, appearance, and adorns. A graphical illustration of the FEI face dataset is represented in Fig. 2.

Also, the ORL database includes 400 facial images, where each subject contains ten different images. For some individuals, the facial images were captured at different lighting variations, periods, facial expressions (not smiling/eye open/smiling/eye closed), and facial details (no glasses/glasses) [20]. A graphical illustration of the ORL data set is presented in Fig. 3.

3.2 Image Pre-processing

The quantization process is carried out to enhance the visibility level of the facial images after collecting the images. The quantization process includes three steps; initially divide the color components into “*n*” and “*p*” shades, then combine red, green, and blue monochromes into a single channel to build the color features, and finally extracts the set of points with the quantized color “*Q*.” Graphically the quantized image is presented in Fig. 4.

Fig. 3 Graphical illustration of ORL dataset



(a) (b)

Fig. 4 **a** Input image, **b** quantized image

3.3 *Feature Extraction*

After denoising the facial images, feature extraction is performed using SURF and MLBP. Compared to other techniques, the selected feature extraction techniques are very simple and efficient in extracting the texture features to achieve better performance in the conditions like illumination condition, lighting variation, facial rotations, etc. Brief explanations about the feature extraction techniques are given below.

3.3.1 Speeded Up Robust Features

The SURF feature is utilized for detecting the blob-like structure when the Hessian matrix determinant is maximized. Consider a point $x = (x, y)$ in a face image I and the hessian matrix $H(x, \sigma)$ at x with scale σ is mathematically denoted in Eq. (1).

$$H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix} \tag{1}$$

where $L_{xx}(x,\sigma), L_{xy}(x,\sigma),$ and $L_{yy}(x,\sigma)$ are indicated as the convolution of Gaussian 2nd order derivation $\frac{\partial^2}{\partial x^2}$ at point x. The scale space is divided into octaves to detect interest points at different scales, where every octave has a series of intervals. The convolution window scale with parameters interval i and octave o is mathematically indicated in Eq. (2). Meanwhile, the relation between the window size and scale σ is denoted in Eq. (3).

$$L = 3 \times \left(i \times 2^o + 1\right) \tag{2}$$

$$L = \sigma \times 9/1.2 \tag{3}$$

Then, the SURF key point is mathematically denoted in Eq. (4).

$$\mathrm{DoH}(x, L) = \max\left(\sum_{k_i=i-1}^{i+1} \sum_{k_x=x-2^o}^{x+2^o} \sum_{k_y=y-2^o}^{y+2^o} \mathrm{DoH}\left(k_x, k_y, o, k_i\right)\right) \geq \lambda \tag{4}$$

where λ is indicated as a positive threshold and DoH is stated as a Hessian matrix determinant. A bright blob centered at (x, y) with scale $L = 3 \times (i \times 2^o + 1)$ is detected if the trace of the hessian matrix is larger than zero.

3.3.2 Modified Local Binary Pattern

The LBP is a productive and effective methodology in image processing applications like face recognition. LBP is a texture feature descriptor, where the central pixel is indicated as g_a. The vector of image pixel p_x is indexed as $g_0, g_1, g_2 \ldots g_{p_{x-1}}$, where the LBP features are obtained by multiplying binomial factor with every binary value, as stated in the Eqs. (5) and (6).

$$\mathrm{LBP}_{px} = \sum_{p_x=0}^{p_x=1} d\left(g_{p_x} - g_a\right) \tag{5}$$

where

$$d(n) = \begin{cases} 1 & n \geq 0 \\ 0 & n < 0 \end{cases} \tag{6}$$

Initially, the squared neighborhood image pixels are estimated in the conventional texture descriptor systems. There will be a variation in rotation invariance if the sign vector varies, so the combination of sign vector and rotation variance is used to extract all the texture features. Therefore, rotation invariance, illumination, and scale are estimated to extract all the texture features, and the magnitude vector in the difference vector is also considered.

Scale Invariance

It is achieved by eliminating the gray value of g_a from circular symmetric neighborhood gray values $g_{p_x}(p_x = 0, 1, \ldots p_{x-1})$, which is mathematically indicated in Eq. (7).

$$T = t\big(g_a, g_0 - g_a, \ldots. g_{p_{x-1}} - g_a\big) \tag{7}$$

where $t(g_a)$ is stated as the luminance value of the facial image. The scale invariance is achieved by considering circular symmetric neighborhood sets by changing the radius $r = 1, 2, 3$ and 4.

Rotation Invariance

The conventional invariant system utilizes only the sign values for facial image texture investigation. In some cases, the sign vector value varies if there are any variations in the image rotations. So, the sign vector along with the magnitude vector is considered to attain rotation invariance, because the magnitude vector remains the same in all the conditions. The mathematical equations to attain rotation invariance are denoted in Eq. (8).

$$\mathrm{LBP}^{id}_{p_{x,r}} = \min\big(\mathrm{RS}\big(\mathrm{LBP}_{P_{x,r}} j\big), \quad j = 0, 1, \ldots P_{x-1}\big) \tag{8}$$

where $\mathrm{RS} = \big(\mathrm{LBP}_{p_{x,r}} j\big)$ performs a bit-wise circle right on x for j times and $\mathrm{LBP}^{id}_{p_{x,r}}$ is denoted as rotation invariant code. The rotation invariance is also utilized to find the illumination changes in the facial images.

Local Difference Sign and Magnitude Transformation

The local difference vector $\big[E_0, E_1, \ldots E_{p_x} - 1\big]$ is proved to be robust for illumination changes by removing g_a. Compared to the input images, there will be more efficient in pattern matching by eliminating g_a from local difference vector $\big[E_0, E_1, \ldots E_{p_x} - 1\big]$. Hence, E_{p_x} is categorized into two elements, as mentioned in the Eqs. (9) and (10).

$$E_{p_x} = S_{p_x} \times M_{p_x} \text{ and } \left\{ \begin{matrix} S_{p_x} = \mathrm{Sign}\left(E_{p_x}\right) \\ M_{p_x} = \left|E_{p_x}\right| \end{matrix} \right\} \tag{9}$$

$$S_{p_x} = \begin{cases} 1 & E_{p_x} \geq 0 \\ -1 & < 0 \end{cases} \tag{10}$$

where M_{p_x} is indicated as E_{p_x} magnitude and S_{p_x} is signified as E_{p_x} sign. Then, the extracted texture features are optimized by GOA to decrease the dimensionality of the extracted feature values.

3.4 Feature Optimization

GOA is a population-based optimization technique that easily handles the unconstrained optimization problems. GOA imitates the behavior of grasshoppers, where the three components (gravity G_r, social relationship S_i, and horizontal wind movement W_i) affect the flying route of grasshopper. In GOA, the searching process is mathematically denoted in Eq. (11).

$$S_i = \sum_{j=1, j \neq i}^{M} s\left(p_{i,j}\right)\widehat{p}_{ij} \tag{11}$$

where p_{ij} is represented as distance between ith and jth grasshopper that is estimated as $p_{ij} = \left|x_j - x_i\right|$, s is stated as strength of social forces and the unit vector from ith to jth grasshopper is indicated as $\widehat{p}_{ij}$, which is mathematically defined in Eq. (12).

$$\widehat{p}_{ij} = \frac{x_j = x_i}{p_{ij}} \tag{12}$$

The function s is the backbone of the social relationship that represents the grasshopper direction in the swarm, and it is mathematically stated in Eq. (13).

$$s(r) = be^{-\frac{r}{l}} - e^{-r} \tag{13}$$

where l is indicated as an attractive length scale, b is represented as an attraction force, r is denoted as the distance between grasshoppers. In the GOA, two types of forces attraction and repulsion are generated. The repulsion force increases when r is in the range of [0, 2.079] that avoids the collision. Correspondingly, the attraction force increases when r is in the range of [2.079, 4] that efficiently handles the swarm cohesion. Equation (14) states the mathematical expression of grasshopper's interaction.

Table 1 Features selected after applying GOA

Sample image	Extracted features	Selected features
1	42 × 3454	42 × 2592
2	42 × 4951	42 × 3028
3	42 × 3885	42 × 3012
4	42 × 4109	42 × 3482
5	42 × 3984	42 × 3091

$$X_i^k = c\left(\sum_{j=1, j\neq i}^{M} c\frac{\mathrm{up}_k - \mathrm{lp}_k}{s} s\left(x_j^d - x_i^d\right)\frac{x_j - x_i}{p_{ij}}\right) + \widehat{T}_k \tag{14}$$

where $\widehat{T}_k$ is indicated as kth dimension in the target, lp_k and up_k are stated as lower and upper bounds, and c is denoted as a coefficient that is utilized for reducing the comfort, repulsion, and attraction regions. The parameter c moves the grasshoppers closer to the target that is estimated as the best solution. The best solution gets updated if the grasshoppers chase the target. The parameter c is estimated using Eq. (15).

$$c = c_{\max} - l\frac{c_{\max} - c_{\min}}{N} \tag{15}$$

where l is indicated as current iteration, N is stated as maximum iterations, $c_{\min} = 0.00001$ and, $c_{\max} = 1$GOA is effective in solving the optimization issues and also the complexity of the algorithm is very low, due to the simple calculation of the distance between the grasshoppers. Table 1 denotes the feature vectors selected after applying GOA. The flow diagram of the GOA is given in Fig. 5.

3.5 Classification

Deep neural network (DNN) is a feed-forward network, where the information is processed by layer-by-layer. In CNN, the layers are stacked where the output of a hidden layer is the input to the succeeding layer. The output of every layer is a function of its internal weights and input. CNN is mathematically defined in Eq. (16).

$$\begin{cases} o_i = X & i = 1 \\ o_i = f_i(z_i) & i > 1 \\ z_i = g_i(o_{i-1}, w_i) \end{cases} \tag{16}$$

where $f_i(.)$ is denoted as the activation function of the ith layer, z_i is stated as weighted operation output of the ith layer, X is indicated as input data, $g_i(.)$ is specified as a weighted operation of the ith layer, w_i is denoted as the weight of the ith layer, and

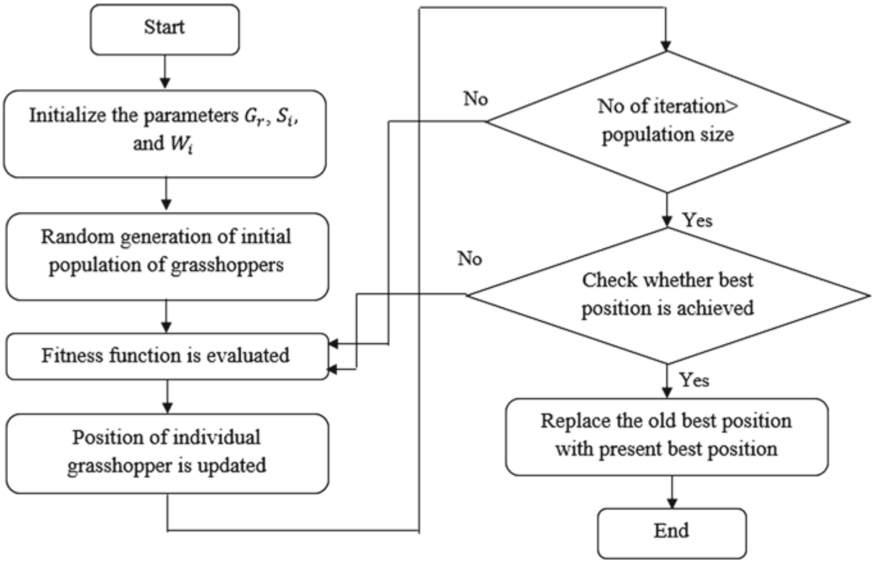

Start
Initialize the parameters G_r, S_i, and W_i
Random generation of initial population of grasshoppers
Fitness function is evaluated
Position of individual grasshopper is updated
No
No of iteration> population size
Yes
No
Check whether best position is achieved
Yes
Replace the old best position with present best position
End

Fig. 5 Flow diagram of GOA

o_i is represented as output of the ith layer. CNN includes three main layers pooling, convolution, and fully connected layers.

The output of the convolutional layer is a convolution operation (Θ) on its weights and inputs and it is known as filters or kernels. There are two kinds of pooling layers average and max-pooling layers that perform a down sampling operation in $r \times c$ window$\left(\mathrm{N}_{r,c}\right)$ for decrease the number of output parameters. In a fully connected layer, the output is a function of the weight multiplied by its input. The CNN layers are mathematically denoted in Eq. (17).

$$\begin{cases} Z_i = w_i \Theta o_{i-1} & \text{if } i\text{th layer is convolutional} \\ z_i = \mathrm{N}_{r,c}\, o_{i-1} & \text{if } i\text{th layer is pool} \\ z_i = w_i o_{i-1} & \text{if } i\text{th the fully connected} \end{cases} \tag{17}$$

The DCNN helps to reduce the error between predicted outputs and training targets. Generally, the minimization of cross-entropy loss is carried out using back propagation and gradient descent in the DCNN classifier. The parameter setting of DCNN is given as follows; convolutional layers are 5, the number of convolution filters in one layer, is 96, the type of activation function is softmax, and pooling window size 5 * 5. The structure of DCNN is given in Fig. 6.

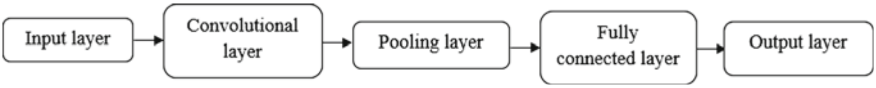


Fig. 6 Structure of DCNN classifier

4 Experimental Investigation

The MATLAB (2019a) environment is used for experimental investigation in a personal computer consist of Intel® Core™ i5-3220 CPU @ 3.30 Hz, 16 GB RAM, and 2 TB hard disc. The performance of the proposed GOA-DCNN model is compared with two existing models IKLDA + PNN [9] and CNN with pre-trained VGG-Face [11] to validate the efficacy of the proposed model in terms of accuracy, precision, f-score, and recall. Though f-measure is used to estimate the test accuracy, and it balances the usage of recall and precision. The mathematical expressions of accuracy, f-score, recall, and precision are presented in the Eqs. (18)–(21).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{FN} + \text{TP} + \text{FP}} \times 100 \tag{18}$$

$$F\text{-score} = \frac{2\text{TP}}{\text{FP} + \text{FN} + 2\text{TP}} \times 100 \tag{19}$$

$$\text{Recall} = \frac{\text{TP}}{\text{FP} + \text{TP}} \times 100 \tag{20}$$

$$\text{Precision} = \frac{\text{TP}}{\text{FP} + \text{TP}} \times 100 \tag{21}$$

where true positive is denoted as TP, true negative is indicated as TN, false positive is stated as FP, and false negative is represented as FN.

4.1 *Quantitative Investigation on FEI Face Dataset*

In this section, the FEI database is taken for validating the performance of the proposed model. The performance of the proposed method is compared with different classification techniques like multiclass support vector machine (MSVM), K-nearest neighbors (KNN), long short-term memory (LSTM), deep belief network (DBN), and DCNN in terms of f-score, accuracy, recall, and precision. In the FEI database, 80% (2240) of the images are used for training and 20% (560) of the images are used for testing, where the collected FEI database facial images are cropped into the size of 32 * 32. By inspecting Table 2, it is concluded that the GOA with DCNN classifier has achieved 98.90% of recognition accuracy, which is better than the other classifiers MSVM, KNN, LSTM, and DBN. Additionally, GOA with DCNN classifier

Table 2 Performance estimation of the proposed model with dissimilar classifiers on FEI face dataset

Method	Precision (%)	Recall (%)	Accuracy (%)	F-score (%)
GOA-KNN	45.89	77.75	80.03	65
GOA-MSVM	67.80	70	86.52	78.92
GOA-LSTM	78.93	89.60	77.90	80.80
GOA-DBN	89.89	93.34	91.17	92.28
GOA-DCNN	**97.50**	**98.78**	**98.90**	**96.75**

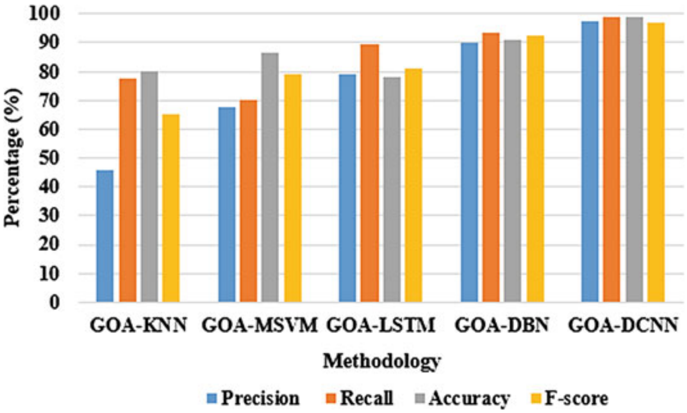


Fig. 7 Graphical depiction of the proposed model with dissimilar classifiers on FEI face dataset

achieved better performance in facial recognition in terms of recall, precision, and f-score. Compared to other classification techniques, GOA with DCNN classifier has faster learning and capacity to manage high dimensional and multivariate data in a dynamic and uncertain environment. Figure 7 presents comparison of performance between proposed and existing models in terms of recall, precision, and f-score on the FEI dataset.

4.2 Quantitative Investigation on ORL Dataset

In Table 3, the performance of the proposed GOA-DCNN model is validated with other classifiers like MSVM, KNN, LSTM, DBN, and DCNN using f-score, accuracy, recall, and precision on the ORL database. The collected ORL dataset images are cropped into the size of 32 * 32, where 80% (320) images are used for training and 20% (40) images are used for testing. From Table 3, it is observed that the proposed GOA-DCNN model has achieved a good performance in face recognition compared to other classification techniques KNN, DBN, MSVM, and LSTM. The

Table 3 Performance estimation of the proposed model with dissimilar classifiers on ORL dataset

Method	Precision (%)	Recall (%)	Accuracy (%)	*F*-score (%)
GOA-KNN	80.74	80.98	87.24	85
GOA-MSVM	88.83	89.30	92.39	90.80
GOA-LSTM	92.64	94.47	91.20	91.09
GOA-DBN	90.09	95.35	94.50	93.30
GOA-DCNN	**97.20**	**97.95**	**99**	**98.61**

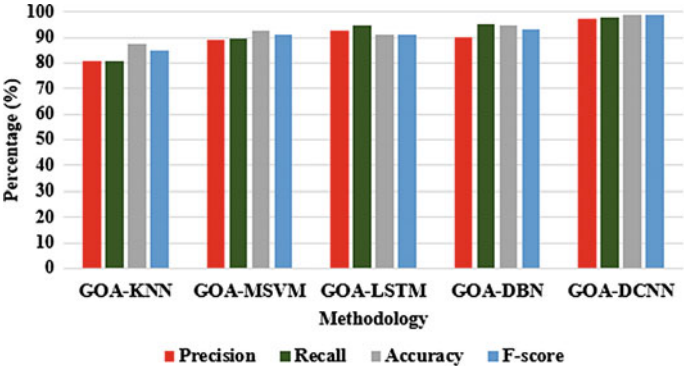


Fig. 8 Graphical depiction of the proposed model with dissimilar classifiers on ORL dataset

proposed GOA-DCNN model has achieved 99% of recognition accuracy which is better compared to other classification techniques. In Table 3, it is observed that the GOA-DCNN classification methodology results in a minimum of 4.50% and a maximum of 11.76% enhanced in recognition accuracy related to other techniques. Respectively, the proposed model GOA-DCNN has attained a good performance in face recognition using f-score, recall, and precision. Figure 8 presents a comparison of performance between proposed and existing models in terms of recall, precision, and f-score on the ORL dataset.

4.3 Comparative Study

Table 4 presents the comparative study of the proposed and existing models. Ouyang et al. [9] developed a hybrid model (IKLDA and PNN) for facial recognition. At first, IKLDA methodology was adapted to lessen the dimension of the extracted feature vectors for retaining the most relevant information about the facial images. Then, the PNN methodology was employed to solve the difficulties of the face recognition system. In this literature, the performance of the developed model was validated on three datasets AR, YALE, and ORL. In the experimental section, the developed

Table 4 Comparative study

Method	Datasets	Accuracy (%)
IKLDA + PNN [9]	ORL	97.22
CNN with pre-trained VGG model [11]	FEI face	90
Proposed model (GOA-DCNN)	ORL	99
	FEI face	98.90

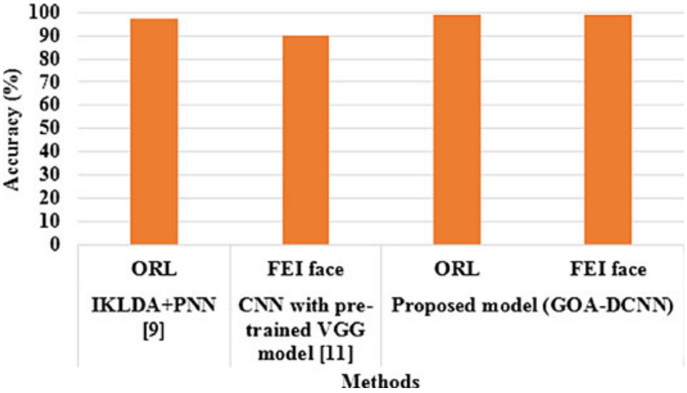


Fig. 9 Graphical comparison of proposed and existing models

model achieved a recognition accuracy of 97.22% in the ORL database. In addition, Elmahmudi and Ugail [11] utilized CNN with a pre-trained VGG model for face recognition. In this paper, the developed model achieved a 90% recognition accuracy in FEI datasets. However, the proposed GOA-DCNN model has achieved better recognition accuracy compared to these existing methods and achieved 1.78%-8.90% of improvement in recognition accuracy. A comparative study is graphically presented in Fig. 9.

5 Conclusion

A new optimization technique with a deep learning classifier (GOA-DCNN) was proposed for facial recognition. The proposed model had three phases feature extraction, optimization, and classification for face recognition. After facial image denoising, MLBP and SURF were applied to extract the features and the extracted feature values are optimized using GOA to decrease the data dimensionality. In the final phase, a DCNN classifier was applied to classify the individual person's facial images. The experimental analysis on ORL and FEI datasets showed that the proposed GOA-DCNN model achieved better performance in face recognition in terms of f-score, accuracy, recall, and precision. Compared to the existing methods, the proposed GOA-DCNN model improved recognition accuracy by up to 1.78–8.90%. In future work, a hybrid optimization technique can be included in the proposed model to further improve the performance of face recognition in the conditions like illumination conditions and light variations.

References

1. Gao G, Yu Y, Yang M, Huang P, Ge Q, Yue D (2020) Multi-scale patch based representation feature learning for low-resolution face recognition. Appl Soft Comput 106183
2. Mi JX, Sun Y, Lu J (2020) Robust supervised sparse representation for face recognition. Cognit Syst Res 62:10–22
3. Zhang G, Porikli F, Sun H, Sun Q, Xia G, Zheng Y (2020) Cost-sensitive joint feature and dictionary learning for face recognition. Neurocomputing 391:177–188
4. Orrù G, Marcialis GL, Roli F (2020) A novel classification-selection approach for the self-updating of template-based face recognition systems. Pattern Recognit 100:107121
5. Shakeel MS, Lam KM (2019) Deep-feature encoding-based discriminative model for age-invariant face recognition. Pattern Recognit 93:442–457
6. Kas M, Ruichek Y, Messoussi R (2018) Mixed neighborhood topology cross decoded patterns for image-based face recognition. Expert Syst Appl 114:119–142
7. Nikan S, Ahmadi M (2018) A modified technique for face recognition under degraded conditions. J Vis Commun Image Rep 55:742–755
8. Gan H (2018) A noise-robust semi-supervised dimensionality reduction method for face recognition. Optik 157:858–865
9. Ouyang A, Liu Y, Pei S, Peng X, He M, Wang Q (2020) A hybrid improved kernel LDA and PNN algorithm for efficient face recognition. Neurocomputing 393:214–222
10. Faraji MR, Qi X (2018) Face recognition under varying illuminations with multi-scale gradient maximum response. Neurocomputing 308:87–100
11. Elmahmudi A, Ugail H (2019) Deep face recognition using imperfect facial data. Future Gener Comput Syst 99:213–225
12. Li H, Suen CY (2016) Robust face recognition based on dynamic rank representation. Pattern Recognit 60:13–24
13. Li Y, Zheng W, Cui Z, Zhang T (2018) Face recognition based on recurrent regression neural network. Neurocomputing 297:50–58
14. Tang J, Su Q, Su B, Fong S, Cao W, Gong X (2020) Parallel ensemble learning of convolutional neural networks and local binary patterns for face recognition. Comput Methods Progr Biomed 105622
15. Dong X, Zhang H, Sun J, Wan W (2017) A two-stage learning approach to face recognition. J Vis Commun Image Rep 43:21–29
16. Roy H, Bhattacharjee D (2018) A novel local wavelet energy mesh pattern (LWEMeP) for heterogeneous face recognition. Image Vis Comput 72:1–13
17. Deng X, Da F, Shao H, Jiang Y (2020) A multi-scale three-dimensional face recognition approach with sparse representation-based classifier and fusion of local covariance descriptors. Comput Electr Eng 85:106700
18. Vijayakumar T (2019) Comparative study of capsule neural network in various applications. J Artif Intell 1(01):19–27
19. Thomaz CE, Giraldi GA (2010) A new ranking method for principal components analysis and its application to face image analysis. Image Vis Comput 28:902–913

20. Jin Z, Yang JY, Hu ZS, Lou Z (2001) Face recognition based on the uncorrelated discriminant transformation. Pattern Recogn 34:1405–1416
21. FEI face dataset. https://fei.edu.br/~cet/facedatabase.html
22. ORL dataset. https://www.cad.zju.edu.cn/home/dengcai/Data/FaceData.html

An Interactive Framework to Compare Multi-criteria Optimization Algorithms: Preliminary Results on NSGA-II and MOPSO

**David F. Dorado-Sevilla, Diego H. Peluffo-Ordóñez,
Leandro L. Lorente-Leyva, Erick P. Herrera-Granda,
and Israel D. Herrera-Granda**

Abstract A problem of multi-criteria optimization, according to its approach, can mean either minimizing or maximizing a group of at least two objective functions to find the best possible set of solutions. There are several methods of multi-criteria optimization, in which the resulting solutions' quality varies depending on the method used and the complexity of the posed problem. A bibliographical review allowed us to notice that the methods derived from the evolutionary computation deliver good results and are commonly used in research works. Although comparative studies among these optimization methods have been found, the conclusions that these offer to the reader do not allow us to define a general rule that determines when one method is better than another. Therefore, the choice of a well-adapted optimization method can be a difficult task for non-experts in the field. To implement a graphical interface that allows non-expert users in multi-objective optimization is proposed to interact and compare the performance of the NSGA-II and MOPSO algorithms. It is chosen qualitatively from a group of five preselected algorithms as members of evolutionary algorithms and swarm intelligence. Therefore, a comparison methodology has been proposed to allow the user for analyzing the graphical and numerical results, which will observe the behavior of algorithms and determine the best suited one according to their needs.

Keywords Evolutionary computation · Multi-objective optimization · Swarm intelligence

---

D. F. Dorado-Sevilla
Universidad de Nariño, Pasto, Colombia

D. H. Peluffo-Ordóñez · L. L. Lorente-Leyva (✉) · E. P. Herrera-Granda · I. D. Herrera-Granda
SDAS Research Group, Ibarra, Ecuador
e-mail: leandro.lorente@sdas-group.com

D. H. Peluffo-Ordóñez
e-mail: dpeluffo@yachaytech.edu.ec

D. H. Peluffo-Ordóñez
Yachay Tech University, Urcuquí, Ecuador

Corporación Universitaria Autónoma de Nariño, Pasto, Colombia

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_5

1 Introduction

Most of the optimization problems that people are commonly facing will have more than one objective simultaneously. In this type of problem, it does not allow one single solution that satisfies all the stated objectives, but rather a set of possible solutions. This set could be very extensive, and if obtaining the best results are desired, then the objective functions must be optimized to find the subset that contains the best solutions. The quality of the obtained set of solutions can vary according to the applied method by taking in count that a general rule which allows defining a method A, as better than a method B, does not exist. In this article, the development of an interactive comparative interface of NSGA-II [1] and MOPSO [2] optimization methods is described, which have been selected, after a review of the state of the art, to represent two of the most used optimization branches: the algorithms inspired by evolutionary theories and those inspired by swarm intelligence. Some applications of these algorithms are proposed for performance optimization and adaptive and intelligent routing of wireless networks using energy optimally [3, 4]. The interface developed in MATLAB allows its user to apply the mentioned algorithms to five different test problems with two objectives and bring the necessary information to conclude which method best suits the user's needs.

This paper is organized as follows: Sect. 2 describes multi-criteria optimization and metaheuristics. Section 3 shows the comparison methodology. Section 4 presents the experimental setup, and Sect. 5 depicts the results and discussion. Finally, the conclusion and the future scope are drawn in Sect. 6.

2 Multi-criteria Optimization

The multi-criteria optimization helps to reach a specific goal, looking for a set of solutions that best adapt to the proposed problem criteria. Depending on the characteristics of the problem, optimizing could involve maximize or minimize the objectives. Thus, a multi-criteria optimization problem in terms of minimization is formally defined as [5].

$$\begin{aligned} &\text{Optimize} \quad y = f(x) = (f_1(x), f_2(x), \ldots, f_k(x)) \\ &\text{s.t.} \quad g(x) = (g_1(x), \ldots, g_m(x)) \leq 0 \end{aligned} \tag{1}$$

where

$$\begin{aligned} x &= (x_1, \ldots, x_n) \in X \subseteq R^n \\ y &= (y_1, \ldots, y_n) \in Y \subseteq R^n \end{aligned}$$

The function $f(x)$ depends of k objective functions, and it can represent real numbers, binary numbers, lists, to-do tasks, etc. The decision vector x contains n

decision variables that identify each solution on the problems space X, which is the set of all the possible elements of the problem. The m restrictions for $g(x)$ limit feasible search areas, where the vector is located x. The objective vector y with k objectives belongs to the objective space Y which is the co-domain of the objective functions. The values, found after solving the objective functions with the decision variables, will be known as functionals.

To classify the best solutions of the solution set, the term dominance is used (Vilfredo Pareto, 1896), which mentions that a Pareto optimum solution is found if it reaches equilibrium, where this solution can't be improved without deteriorating another. Formally, since u and v are vectors contained in the decision space $f(u)$ and $f(v)$ then corresponding functionals, it can be said in minimization terms that:

The dominant vector will be which has the minor *functional*. Then,

$$u \prec v (u \text{ dominates a } v) \quad \text{if and only if} \quad f(u) < f(v).$$
$$v \prec u (v \text{ dominates a } u) \quad \text{if and only if} \quad f(v) < f(u).$$

Solutions are not compatible if none of the vectors dominates each other. This is:

$$u \sim v (u \text{y } v \text{ are not comparable}) \quad \text{if and only if} \quad f(u) \neq f(v) \wedge f(v) \neq f(u).$$

The optimization methods try to find, in the decision space, the set called Pareto optimum defined as $X_{\text{true}} = \{x \in X | x \text{ is not dominated respect a } X\}$, for succeeding, reaching the Pareto front in the objective space defined as $Y_{\text{true}} = F(X_{\text{true}})$ [6].

2.1 Metaheuristics

They are algorithms that modify variables trough time, guided by expert knowledge through the feasible area of the decision space in an iterative manner. The best results are obtained by applying improvements to a set of initial solutions, based on the mentioned concept of dominance, to discard the least suitable solutions [7].

2.1.1 NSGA-II (Non-dominated Sorting Genetic Algorithm)

It is a genetic algorithm chosen to represent evolutionary algorithms [8]. It is widely used in the literature for solving multi-criteria optimization problems, as shown in [9, 10]. It is considered one of the best methods for its strategies to maintain elitism and diversity in the search for optimal solutions, using Darwinian natural selection analogy, which establishes that only the fittest individuals survive and reproduce to generate a new generation with improved aspects. In Algorithm 1 is detailed the pseudocode of the algorithm proposed in [1].

Algorithm 1 Algorithm NSGA-II

```
Start a population:
  Generate an aleatory population $P$.
  Evaluate aptitude.
  Assign a level based on Pareto dominance - "arrange".
  Generate a $P$ population as follows:
    Selection by binary tournament.
    Recombination and mutation.
  For $i = 1$ up to generation number Do:
    For Father and son population Do:
          Assign a level based on Pareto dominance and arrange.
      Generate the set of not dominated fronts.
        Sum solutions to the next generation, starting by the hierarchically first and use the stacking factor (crowding) on each front.
    End For.
     Select points on the lowest front, which are out of the stacking factor distance.
    Create the next generation:
      Select by binary tournament.
      Recombination and mutation.
End For.
```

Initially, the algorithm randomly creates an initial population of feasible P_0 solutions of N size and then forms a Q_0 population also of N size using binary tournament selection, recombination, and mutation. The next step is to combine the two populations in such a way that from the new population $R_0 = P_0 + Q_0$, using selection, mutation, and recombination, a new P_1 population is born. The process is repeated in the following generations as shown in Fig. 1.

2.1.2 Multi-objective Particle Swarm Optimization (MOPSO)

It is an algorithm representative of swarm intelligence, popular in the literature for solving multi-criteria optimization problems, due to its good performance as can be seen [11, 12]. Besides, it is a collective intelligence algorithm with similar search behavior to the flight of starlings. These birds move in flocks and coordinates the direction and speed of their flights, so that a subgroup of the population, in response to an external stimulation, transmits clearly and immediately the state of their movement to the rest of the group. Each individual maintains a maximum susceptibility to any change in the flight of their neighbors, who react quickly to external stimulation and transmit the information to the whole group [13]. The pseudocode proposed in [2] is shown in the Algorithm 2.

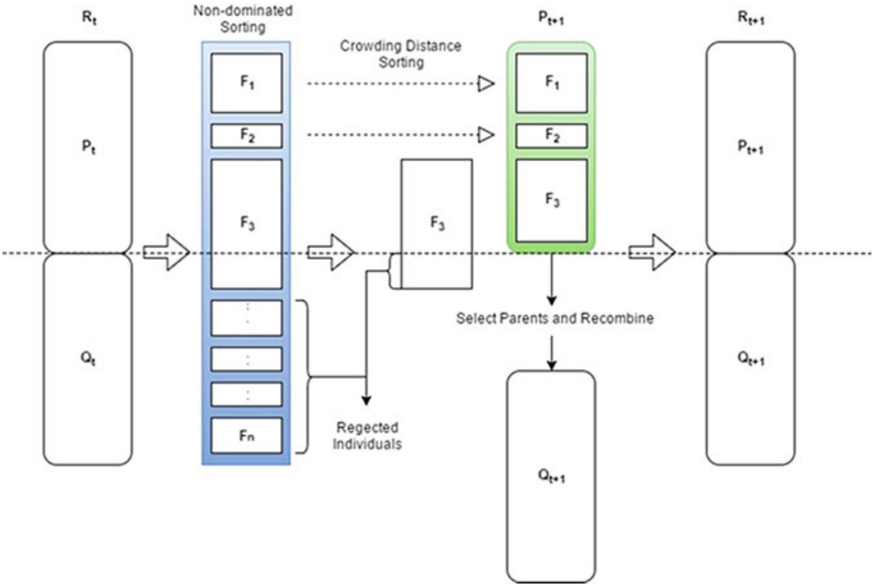


Fig. 1 NSGA-II algorithm search process

Algorithm 2 Algorithm MOPSO

```
Initialization phase:
For $i$ = 1 to $N$ Do:
    Apply to $position_i$ an aleatory value in the rank [$X_{min}$,$X_{max}$]
    Assign zeros to $velocity_i$.
    Evaluate $functional_i$.
    $Best\_position_i$ equal to $position_i$.
    Best_$functional_i$ equal to $funcional_i$.
End For
Evaluate dominance.
Save not dominated individuals on a repository and assign a quadrille index.
Search phase:
For $i_t$ = 1 to the maximum number of iterations Do
    For $i$ = 0 to $N$ Do.
        Chose the global best from repository.
        Calculate $velocity_i$.
        Calculate $position_i$.
        Keep individuals within the search space and evaluate $functional_i$.
        Apply the mutation factor.
        If $position_i$ dominates the $best\_position_i$ in its memory then: $Pbest_i$ = $position_i$.
    End For.
    Evaluate dominance.
    Update repository and evaluate dominance.
    Eliminate the dominated individuals from repository.
    Decrease the factor w.
End For
```

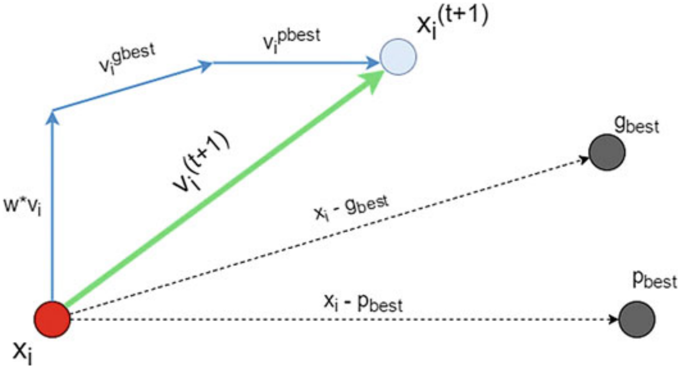


Fig. 2 Individual's change of position

The MOPSO performs the search for optimal solutions imitating the behavior of a flock in search of food. The position of each individual is obtained from the following equations.

$$\begin{aligned} v_{id}(t+1) = w * v_{id}(t) + c_1 * r_1 * [pbest_{id} - x_{id}(t)] \\ + c_2 * r_2 * [(gbest_{id} - x_{id}(t)] \end{aligned} \tag{2}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \tag{3}$$

where v_{id} is the speed value of individual i in the d dimension; $c1$ is the cognitive learning value; $c2$ is the global learning factor; $r1$ and $r2$ are random values uniformly distributed in the range [0.1]; x_{id} is the position of individual i in the d dimension; *pbestid* is the value in the d dimension of the individual with the best position found by individual i; and *gbestd* is the value in the d dimension of the individual in the population with the best position. The value w is important for the convergence of the algorithm. It is suggested that $c1$ and $c2$ take values in the range [1.5, 2] and w in the range [0.1, 0.5] [14]. The change of position is shown in Fig. 2.

3 Comparison Methodology

It is necessary to establish guidelines that allow understanding of how the two methods of optimization perform against certain objective functions. To evaluate its performance, four metrics are used to measure the convergence to the optimum Pareto front [15, 16].

*3.1 Error Ratio (*E*)*

This measure determines the portion of individuals in the set of solutions found by the algorithm Y_{known} that belongs to the Pareto optimal solution Y_{true}, where a value of $E = 0$ is ideal. Formally, it is defined as follows:

$$E \triangleq \frac{\sum_{i=1}^{N} e_i}{N} \tag{4}$$

$$e_i = \begin{cases} 0, & \text{if a vector of } Y_{\text{known}} \text{ is in } Y_{\text{true}} \\ 1, & \text{otherwise} \end{cases} \tag{5}$$

*3.2 Generational Distance (*DG*)*

This measure determines the solutions which are found by the Pareto optimal algorithm. Mathematically, it is defined as:

$$\text{DG} = \sqrt{\frac{\sum_{i=1}^{N} d_i^2}{N}} \tag{6}$$

where d_i is the Euclidean distance between each objective vector that belongs to the solution set found and its closest corresponding member in the real optimal Pareto front.

*3.3 Spacing (*S*)*

Verifies the dispersion of the elements of the Pareto set X found by the algorithm. Knowing the individuals at the extremes of the set, this measure proposes to use the variance of the distance between neighboring vectors of the current X set.

$$S \triangleq \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\overline{d} - d_i\right)^2} \tag{7}$$

For two objective functions, $d_i = \min_j\left(\left|f_1^i(x) - f_1^j(x)\right| + \left|f_2^i(x) - f_2^j(x)\right|\right)$ is the Euclidean distance between consecutive solutions of Y_{known} $i, j = 1, 2, \ldots, n$, where n is the number of individuals in the set.

4 Experimental Setup

To test the performance of the optimization algorithms, in [17] the test functions, proposed by Zitzler, Deb, and Thiele, are used. The functions ZDT1, ZDT2, ZDT3, ZDT4, and ZDT6 allow analyzing the behavior of the algorithms when optimizing five different Pareto fronts. The optimal fronts of the five functions are given for $g(x) = 1$.

4.1 ZDT1 Function

This function has a convex and continuous front. With $n = 30$ as the number of decision variables and x_i in the [0, 1] rank.

$$f_1(x) = x_1 \tag{8.}$$

$$g(x) = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i,$$

$$h(f_1, g) = 1 - \sqrt{\frac{f_1}{g}},$$

$$f_2 = g(x) * h(f_1(x), g(x)) \tag{9}$$

4.2 ZDT2 Function

This function has a convex and continuous Pareto front. With $n = 30$ as the number of decision variables and x_i in the [0, 1] rank.

$$f_1(x) = x_1, \tag{10}$$

$$g(x) = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i,$$

$$h(f_1, g) = 1 - \left(\frac{f_1}{g}\right)^2,$$

$$f_2 = g(x) * h(f_1(x), g(x)) \tag{11}$$

4.3 ZDT3 Function

This function has a discontinuous Pareto front segmented into five parts. With $n = 30$ as the number of decision variables and x_i in the [0, 1] rank.

$$f_1(x) = x_1 \tag{12.}$$

$$g(x) = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i,$$

$$h(f_1, g) = 1 - \sqrt{\frac{f_1}{g}} - \frac{f_1}{g}\sin(10\pi f_1),$$

$$f_2 = g(x) * h(f_1(x), g(x)) \tag{13}$$

4.4 ZDT4 Function

This is a multi-modal function that has several convex and continuous Pareto fronts. With $n = 10$ as the number of decision variables, x_i in the [0, 1] rank and x_i in the $[-5, 5]$ rank for $i = 2, \ldots, n$.

$$f_1(x) = x_1 \tag{14}$$

$$g(x) = 1 + 10(n-1) + \sum_{i=2}^{n}\bigl(x_i^2 - 10\cos(4\pi x_i)\bigr),$$

$$h(f_1, g) = 1 - \sqrt{\frac{f_1}{g}},$$

$$f_2 = g(x) * h(f_1(x), g(x)) \tag{15}$$

4.5 ZDT6 Function

This function has a non-convex and continuous Pareto front. With $n = 10$ as the number of decision variables and x_i in the [0, 1] rank.

$$f_1(x) = 1 - \exp(-4x_1) * \sin^6(6\pi x_1) \tag{16}$$

$$g(x) = 1 + 9\left[\frac{\sum_{i=2}^{n} x_i}{9}\right]^{0.125}$$

$$h(f_1, g) = 1 - \left(\frac{f_1}{g}\right)^2$$

$$f_2(x) = g(x) * h(f_1(x), g(x)) \tag{17}$$

4.6 Parameters

Tables 1 and 2 show the parameters used in the execution and simulation of NSGA-II and MOPSO algorithms.

Table 1 Parameters for execution of the NSGA-II algorithm

Parameters	
Population size	100
m	30
Number of iterations	200
Range, decision variables	[0 1]
Crossover rate	0.8
Mutation rate	0.033
Number of mutants	20

Table 2 Parameters for execution of the MOPSO algorithm

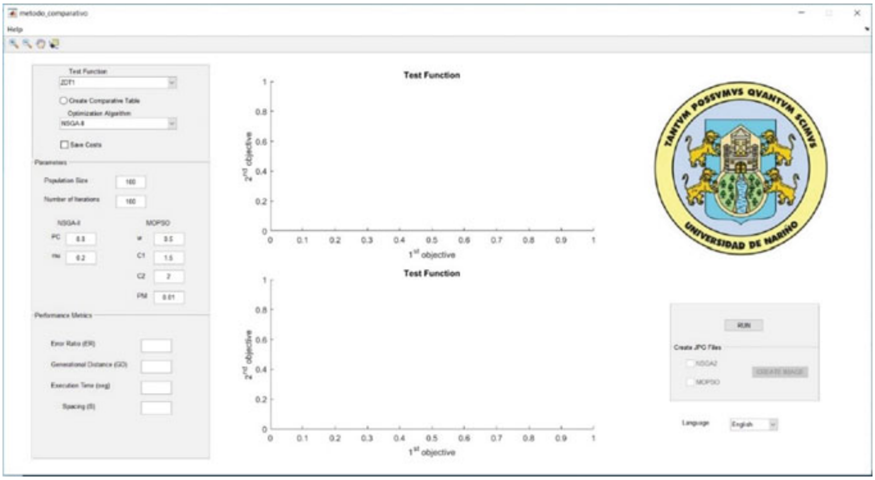
Parameters	
Population size	100
Decision variables	30
Number of iterations	200
Range, decision variables	[0 1]
w	0.5
Wdamp	0.99
Mutation rate	0.01
$c1$	1
$c2$	2

5 Results and Discussion

A graphic interface in Fig. 3 using MATLAB was developed, in which the user is allowed to apply the NSGA-II and MOPSO algorithms to the five ZDT test functions mentioned above, to obtain numerical results of the proposed performance measures. In addition, this interface shows iteratively how each algorithm tracks the best possible solutions in the search space. In order to execute the interface, it is necessary to introduce certain evaluation parameters that guide the search of each algorithm. These parameters are loaded for each test function automatically. In the following Tables 1 and 2, the parameters loaded in the interface for the two algorithms, and their respective test functions are shown.

Since the NSGA-II algorithm is based on a population for the solutions search, the N size of this population must be defined. A stop parameter is needed to stop the search, in this case, a maximum of *MaxIt* iterations. To create the population, define the number of *parents* to generate a group of *descendants*, where Pc is the crossing rate. The number of mutants is defined as $\mathrm{nm} = \mathrm{round}(\mathrm{pm1} * N)$, where Pm1 is the mutation rate. Table 3 shows the parameters used in the NSGA-II to evaluate the five defined test functions.

Like the previous algorithm in the MOPSO, you must define the N size of the individuals that will take flight in search of optimal solutions and a *MaxIt* stop parameter. As for the search procedure, the change of individual's position is fundamental, the parameters w, $c1$, $c2$ defined in Eq. (2) must be defined. To generate diversity,



Test Function: ZDT1
Optimization Algorithm: NSGA-II
Population Size: 100
Number of Iterations: 100
NSGA-II
PC: 0.8
mu: 0.2
MOPSO
w: 0.5
C1: 1.5
C2: 2
PM: 0.01
Language: English

Fig. 3 Interactive comparator developed in a graphical interface using MATLAB. More information and MATLAB scripts available at: https://sites.google.com/site/degreethesisdiegopeluffo/interactive-comparator

Table 3 Evaluation parameters NSGA-II

Parameters	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
N	100	100	100	100	100
MaxIt	500	500	500	500	500
Pc	0.67	0.63	0.63	0.67	0.67
Pm1	0.33	0.33	0.33	0.33	0.33

Table 4 Evaluation parameters MOPSO

Parameters	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
N	100	100	100	100	100
MaxIt	100	100	100	100	200
W	0.5	0.5	0.5	0.5	0.5
$c1$	1.5	1.5	1.5	1.5	1.5
$c2$	2	2	2.5	2	2.5
Pm2	0.1	0.1	0.5	0.1	0.5

the algorithm simulates turbulence in flight using a mutation operator. In each iteration, all individuals are assigned a mutation probability (Pm2). Table 4 shows the parameters used to evaluate the five defined test functions.

To allow the user to conclude the results easily, the "Create Comparative Table" function is created in the interface, which executes automatically each algorithm ten times in a row and creates an excel file that contains a table with the numerical results of the performance measures for each execution.

5.1 ZDT1 Results

In Table 5, the results obtained with the ZDT1 execution are presented, where it can be evidenced that the performance of the MOPSO when optimizing a problem with continuous and convex front is better than the NSGA-II, where the E and DG metrics are very close to the real Pareto front. It is also noted that the swarm intelligence algorithm is much faster. It is also shown that according to the S metric, the NSGA-II has a better dispersion than the MOPSO.

Figures 4 and 5 show the Pareto front of the ZDT1 function and the solutions distribution.

In the previous figures, it is shown that the analysis made in the execution of the ZDT1 function with the developed interface, where a user will be able to make in the interface, the analysis for the rest of the presented problems (ZDT2, ZDT3, ZDT4, and ZDT6). And obtain in this way, the behavior of each algorithm is used against

Table 5 ZDT1 results

Execution	E_NSGA	E_MOPSO	DG_NSGA	DG_MOPSO	Time NSGA	Time MOPSO	S_NSGA	S_NSGA
1	1.000	0.000	0.012	0.000	318.006	44.153	0.007	0.035
2	0.990	0.000	0.013	0.001	384.542	42.192	0.006	0.023
3	0.970	0.000	0.016	0.001	466.853	40.592	0.009	0.023
4	1.000	0.000	0.013	0.001	517.460	40.701	0.009	0.018
5	1.000	0.000	0.021	0.001	574.547	41.070	0.016	0.022
6	1.000	0.000	0.013	0.000	645.206	42.438	0.007	0.020
7	1.000	0.000	0.018	0.001	712.545	41.705	0.013	0.021
8	0.990	0.000	0.013	0.000	777.596	40.710	0.007	0.021
9	0.970	0.000	0.013	0.001	893.973	42.343	0.007	0.018
10	1.000	0.000	0.015	0.001	1196.094	39.570	0.008	0.019
Average	**0.992**	**0.000**	**0.015**	**0.001**	**648.682**	**41.547**	**0.009**	**0.022**

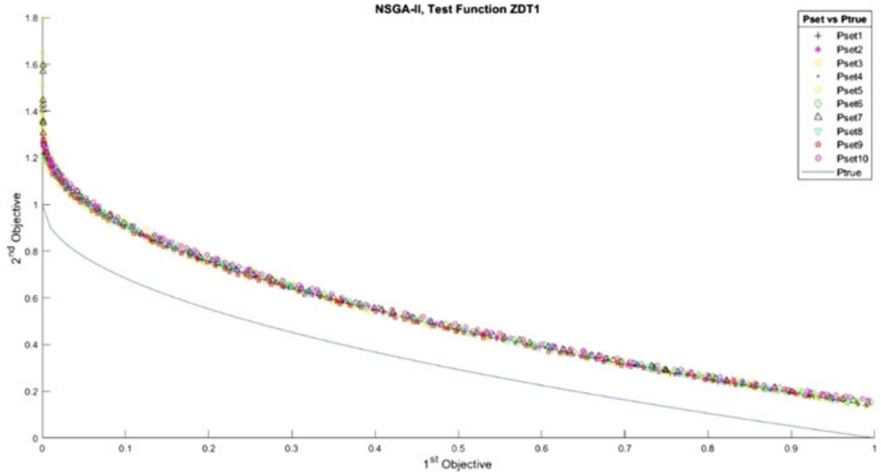


Fig. 4 NSGA-II solutions with ZDT1 optimization versus continuous convex Pareto front

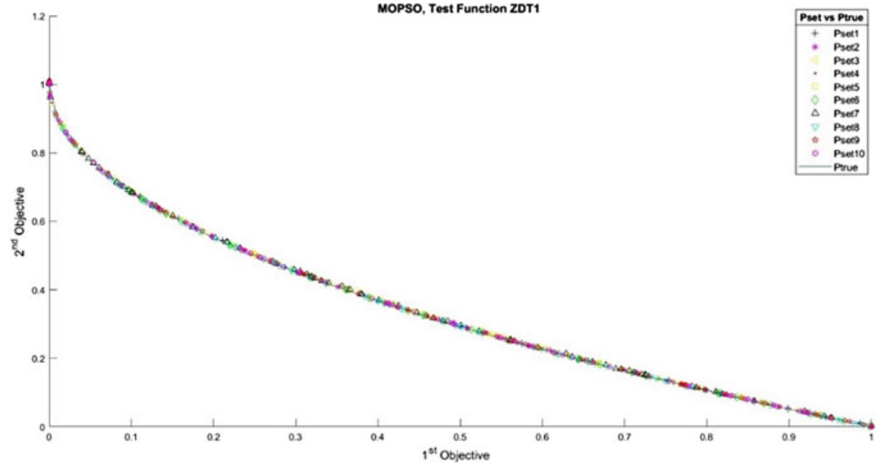


Fig. 5 MOPSO solutions with ZDT1 optimization versus continuous convex Pareto front

all the given conditions. Determine the algorithm that has the best performance and obtains the best results.

6 Conclusion and Future Scope

The results thrown by an optimization algorithm can reach different quality levels, depending on the variation of the evaluation parameters. Therefore, the comparative studies of multi-criteria optimization methods available in the literature, limit the reader performance analysis of the algorithms, by basing their experiments on fixed evaluation parameters.

The development of the interactive comparative interface offers the possibility of easily carrying out an optimization process in an intuitive way. This interactive interface allows the user to choose the optimization algorithm and the test function that he wants to optimize according to the Pareto front of interest and establish a man–machine communication, through several inputs defined as parameters of evaluation, which can be modified to obtain a dynamic graphic and numerical response.

Using the mapping of the objective functions, the set of solutions found in the target space at the end of each iteration can be observed, allowing them to observe the search procedure in a dynamic way. The user can accurately measure the performance of the algorithms by evaluating the numerical results of the performance measures, which are on the final set of solutions found. For the above, a not necessarily expert user will have a greater understanding of the optimization process and will choose the appropriate method more easily, according to their needs. As future work, it is proposed to expand the number of optimization algorithms and add new test functions such as three objective functions and so on.

Acknowledgements The authors are greatly grateful for the support given by the SDAS Research Group (https://sdas-group.com/).

References

1. Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6(2):182–197
2. Coello Coello C, Lechuga M (2002) MOPSO: a proposal for multiple objective particle swarm optimization. In: Proceedings of the 2002 Congress on evolutionary computation, CEC'02, pp 1051–1056
3. Rahimunnisa K (2019) Hybridized genetic-simulated annealing algorithm for performance optimization in wireless Adhoc network. J Soft Comput Paradigm 1(01):1–13
4. Shakya S, Pulchowk LN (2020) Intelligent and adaptive multi-objective optimization in WANET using bio inspired algorithms. J Soft Comput Paradigm 2(01):13–23
5. Deb K, Agrawal S, Pratap A, Meyarivan T (2000) A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: Nsga-II. In: International conference on parallel problem solving from nature. Springer, pp 849–858
6. Veldhuizen DAV, Lamont GB (2000) Multiobjective evolutionary algorithms: analyzing the state-of-the-art. Evolut Comput 8(2):125–147
7. Melián B, Pérez JAM, Vega JMM (2003) Metaheurísticas: Una visión global. Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial 7(19)
8. Kannan S, Baskar S, McCalley JD, Murugan P (2009) Application of NSGA-II algorithm to generation expansion planning. IEEE Trans Power Syst 24(1):454–461
9. Kwong WY, Zhang PY, Romero D, Moran J, Morgenroth M, Amon C (2014) Multi-objective wind farm layout optimization considering energy generation and noise propagation with Nsga-II. J Mech Des 136(9):091010
10. Lorente-Leyva LL et al (2019) Optimization of the master production scheduling in a textile industry using genetic algorithm. In: Pérez García H, Sánchez González L, Castejón Limas M,

Quintián Pardo H, Corchado Rodríguez E (eds) HAIS 2019. LNCS 11734, Springer, Cham, pp 674–685

11. Robles-Rodriguez C, Bideaux C, Guillouet S, Gorret N, Roux G, Molina-Jouve C, Aceves-Lara CA (2016) Multi-objective particle swarm optimization (MOPSO) of lipid accumulation in fed-batch cultures. In: 2016 24th Mediterranean conference on control and automation (MED). IEEE, pp 979–984
12. Borhanazad H, Mekhilef S, Ganapathy VG, Modiri-Delshad M, Mirtaheri A (2014) Optimization of micro-grid system using MOPSO. Renew Energy 71:295–306
13. Marro J (2011) Los estorninos de san lorenzo, o cómo mejorar la eficacia del grupo. Revista Española De Física 25(2):62–64
14. Parsopoulos KE, Vrahatis MN (2002) Recent approaches to global optimization problems through particle swarm optimization. Nat Comput 1:235–306
15. Van Veldhuizen DA, Lamont GB (1999) Multiobjective evolutionary algorithm test suites. In: Proceedings of the 1999 ACM symposium on applied computing. ACM, pp 351–357
16. Eberhart R, Kennedy J (1995) A new optimizer using particle swarm theory. In: Micro machine and human science. In: Proceedings of the Sixth International Symposium on MHS'95. IEEE, pp 39–43
17. Zitzler E, Deb K, Thiele L (2000) Comparison of multi-objective evolutionary algorithms: empirical results. Evolut Comput 8:173

A Dynamic Programming Approach for Power Curtailment Decision Making on PV Systems

Yasmany Fernández-Fernández, Leandro L. Lorente-Leyva, Diego H. Peluffo-Ordóñez, and Elia N. Cabrera Álvarez

Abstract The new grid codes for large-scale photovoltaic power plants require power curtailment despite the variation of solar irradiance. This power curtailment is been developed considering one reference of active power. However, this value is chosen according to the demand, but it is not considering other aspects as solar irradiance or cloudiness. Therefore, this article presents a novel approach to tackle this issue. For this, stochastic dynamic programming is considered to optimize the decision of the power reference every hour considering the solar irradiance and cloudiness during different stages of the day. The results obtained are compared with the performance of the photovoltaic power plant, and it is a referential approach that uses the maximum power point tracking algorithms for the construction of referential power intervals over longer time intervals.

Keywords Maximum power point tracker (MPPT) · Photovoltaic power plant (PVPP) · LS-PVPP

Y. Fernández-Fernández
Universidad Politécnica Estatal del Carchi, Tulcán, Ecuador
e-mail: yasmany.fernandez@upec.edu.ec

Y. Fernández-Fernández · L. L. Lorente-Leyva (✉) · D. H. Peluffo-Ordóñez
SDAS Research Group, Ibarra, Ecuador
e-mail: leandro.lorente@sdas-group.com

D. H. Peluffo-Ordóñez
Corporación Universitaria Autónoma de Nariño, Pasto, Colombia

Yachay Tech University, Urcuquí, Ecuador
e-mail: dpeluffo@yachaytech.edu.ec

E. N. C. Álvarez
Universidad de Cienfuegos, Cienfuegos, Cuba
e-mail: elita@ucf.edu.cu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_6

1 Introduction

The photovoltaic power plants of the LS-PVPP type operation represent a problem for operators. Normally, there is a variability in the behavior of solar irradiance mainly during the day. Using photovoltaic inverters, it has been possible to integrate voltage and frequency support and active and reactive power control [1].

One of the issues to solve is to find an adequate value of the active power that the LS-PVPP must supply to comply with the grid codes despite the variation of solar irradiance. The reduction of the active power to a fixed power is called by the grid operators as "power curtailment." Currently, this curtailment is only performed due to demand and grid behavior. The decision of this value is performed by the grid operator, and the LS-PVPP must supply this power at any moment of the day. However, this decision is not developed considering solar irradiance variability, temperature, costs or any other factor that could affect directly to the operation of the LS-PVPP.

For the management of active power in renewable energy, some optimization techniques have been used [2, 3]. One of these is the stochastic dynamic programming as it helps to add uncertainty scenarios due to the variability of the input energy [4]. For instance, Papavasiliou et al. [5] use this technique to optimize the curtailment of a wind power plant according to the grid response, the technical requirements and the variation of wind power. The challenges of the stochastic dynamic programming addressed by this research are (i) the appropriate selection of the weighing scenarios and (ii) the computational intractability of the resulting problem.

In the photovoltaic field, this optimization technique has not been used for power curtailment as it is a new requirement asked by the grid codes as the case of Puerto Rico [6]. Thus, the aim of this work is to find the optimal power point that the LS-PVPP has to supply considering solar irradiance, cloudiness and in hourly basis using a stochastic dynamic programming approach.

The rest of the paper is structured as follows: Sect. 2 presents an explanation of stochastic programming. Section 3 shows the formulation of the problem considering the uncertainty of cloudiness and solar irradiance during different stages of the day. In Sect. 4, a scenario is tested for the given algorithm and the results are presented. Finally, the conclusion and the future scope are described in Sect. 5.

2 Background and Problem Statement

Linear models represent the basis for formulating linear problems, in general, a basic linear model has the following form:

$$\begin{aligned} &\min_x \mathbf{c}^T \mathbf{x} \\ &\text{st} \quad \mathbf{A}\mathbf{x} = \mathbf{b} \\ &\qquad \mathbf{x} > \mathbf{o_n}, \end{aligned} \tag{1}$$

where $\mathbf{x}$, $\mathbf{c}$, $\mathbf{b} \in \mathrm{R}^{\mathbf{n}}$, $\mathbf{A} \in \mathrm{R}^{\mathbf{nxn}}$, and $0_{\mathbf{n}}$ is a n-dimensional all-zeros vector.

To make a forecast, it is necessary to have quantitative information on-demand behavior over time, with analysis using classic statistical techniques such as ARIMA, Holt-Winters, among others, the most widely used to predict their behavior [7, 8]. Through this representation, thousands of people doing research have been able to represent their problems through a standard mathematical model. The modeling of dynamic programming problems has a considerable focus on complexity with respect to linear problems [9].

The challenge of dynamic programming is decision making. Two general procedures in the deterministic sampling or stochastic sampling approaches are important for considering the problem to solve. The first approach involves the representation of the stochastic process through a decision tree from which an associated equivalent deterministic problem is obtained that is solved by an optimization technique that may or may not lead to a discretization of the problem with a deterministic approach. The second approach assumes the complete tree of the problem which cannot be enumerated, so it is necessary to approximate successive sampling that according to the bibliography can be done with two approaches of exterior and interior sampling [7].

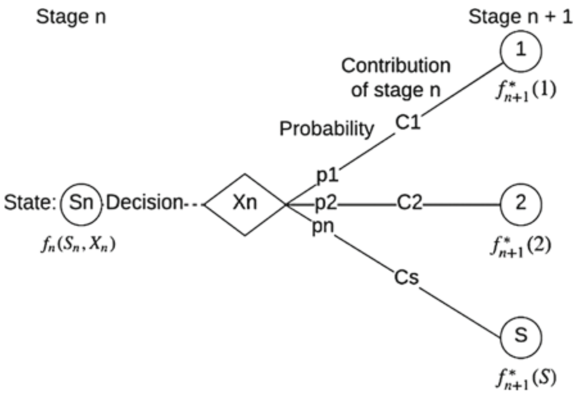
Another approach is presented in [10], which combines clustering algorithms with dynamic programming for designing a demand response strategy applied to residential electric water heaters. In [11], authors develop a multi-objective optimization algorithm to manage energy in radio networks and reduce consumption by minimizing communication delays. Other researchers [12] use some methods to optimize the performance of a wireless network to retain the energy level of the devices.

In [8], authors use a real-time dynamic economic load dispatch integrated with renewable energy curtailment to detect a minimum amount of supply–demand mismatch in advance, managing reliably for a considerable time horizon. One last experience using dynamic models is presented in [13] where the authors propose an improved multistage converter topology intended for single-phase solar rooftop applications with battery storage.

Stochastic dynamic programming refers to the existence of a probability to obtain the results of a near state.

By expanding the relationships represented in Fig. 1 to include all states and possible decisions in all stages, one gets what is often called the decision tree, which, if not very large, provides a useful way to summarize the different possibilities. Let us define $f_n(S_n, x_n)$ and $f^*_{n+1}(S_{n+1}, x_{n+1})$, respectively, as the instantaneous and optimal objective function value in terms of the state S and the independent variable x. Due to the probabilistic scheme, the relationship between $f_n(S_n, x_n)$ and $f^*_{n+1}(S_{n+1}, x_{n+1})$ is more complicated than the deterministic case, and it will depend on the general form of the objective function.

Fig. 1 Probabilistic dynamic problem scheme



3 Formulation of the Proposed Mathematical Model

PV is used as a basic unit for large-scale photovoltaic power plants. This generator controls the active power at every instant according to the solar irradiance. Commonly, the PV generator follows the maximum power possible to get at the given conditions of solar irradiance and ambient temperature. The control is called the maximum power point tracker (MPPT). Different algorithms have been developed to track this point at any instant [14].

Because grid code requirements are not necessary to track the maximum power point at each instant, a referential power can be used. To optimize this power reference considering solar irradiance and cloudiness, the approach of this model is based on a decision tree of stochastic dynamic programming.

For this model, the day will be divided into several parts that are the stages of the model (i). Each stage i represents a constant time in which the system calculates the MPPT, if the MPPT is calculated every 5 min, then $i \in \{1, \ldots, n\}$ corresponds to 5, …, nt minutes where the system collects MPPT data, as shown in Eq. (2).

$$i = n \rightarrow nt \text{ minutes} \tag{2}$$

The range of active power values that can be signed in the stage i can be given as follows:

$$\delta_i\left[\theta_i, \theta_{i+1}\right] \tag{3}$$

where θ_i the maximum power point in stage i and θ_{i+1} is the maximum power point in stage $i + 1$. Considering Eq. (3), the optimal power function in stage i (ρ_i) can be defined depending on the range of active power and also on the optimized value of active power for the next stage (θ_{i+1}).

$$\rho_i(\delta_i)\rho(\delta_i, \theta_{i+1}) \tag{4}$$

where δ_i, θ_i is the contribution to the objective function of stages i, $i + 1$, …, n. If the system is in the state δ_i in stage i, the immediate decision is θ_i and from now on optimal decisions are taken in the form:

$$\rho_i^*(\delta_i) = \max_{\rho} \rho(\delta_i, \theta_i)$$
$$\text{s.t.} \quad \theta_i \geq 0 \tag{5}$$

Then, the average accumulated contribution of power θ_i up to stage $i = n$ is expressed as follows:

$$C(\theta_i) = \frac{1}{n}\sum_{i=1}^{n}\theta_i \tag{6}$$

The decision of the maximum value possible between the current power δ_i and power θ_{i+1} can be expressed as follows:

$$\rho_i^*(\delta_i, \theta_i) = \sup\{[\delta_i, \theta_{i+1}]\} \tag{7}$$

So, the optimal decision for each stage i is written in the form:

$$\rho_i^{**}(\delta_i) = \max_{\rho*}\{kb_i(C(\theta_i) + \rho_i^*(\delta_i, \theta_i))\} \tag{8}$$

where kb_i is the clarity index in stage i. This index is calculated by daily accumulated values in two temporary resolutions, hourly and daily. The time series are grouped in ranges of $kb = 0.2$ assuming that the behavior of the fluctuations of the radiation for each range is very similar. Table 1 shows this classification based on the value of the clarity index [15].

The result of this decision is a recursive function, Eq. (8), which considers the value of the accumulated referential power that follows each stage. This model is considering the possible maximum power of the solar irradiance at each stage and the optimal possible power considering different sky conditions. The recursive form of this function is illustrated in Fig. 2.

Table 1 Clarity index for different sky types

kb values	Type of sky
$kb \leq 0.2$	Completely covered
$0.2 \geq kb \leq 0.4$	Mostly covered
$0.4 \geq kb \leq 0.6$	Partially covered
$0.6 \geq kb \leq 0.67$	Mostly clear
$kb \geq 0.67$	Completely clear

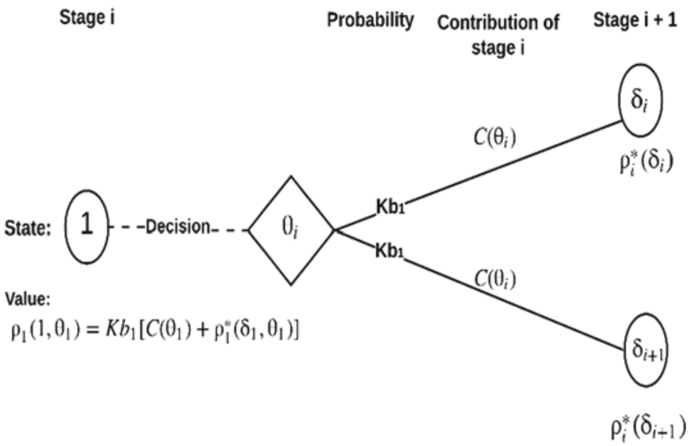


Fig. 2 Representation of the problem through a stochastic dynamic programming model

4 Results and Discussion

To study the model proposed, real data from a photovoltaic power plant located in the north of Ecuador is considered. Figure 3 shows the maximum power supplied by the PVPP at different solar irradiance taken each 5 min for a single day in 2018.

The proposed model and the data from the PVPP are introduced in the software "Wolfram Mathematica 11.2." To simulate the model, the day (from 6h30 to 18h30)

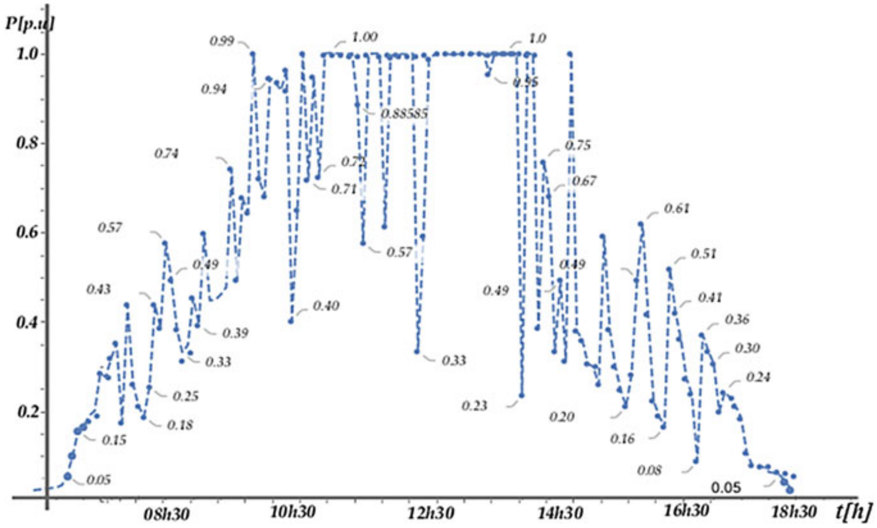


Fig. 3 Active power supplied by a photovoltaic power plant located at the north of Ecuador

is divided into six parts of 2 h each. Figure 3 shows the representation of the data of a common day used in the proposed simulation. Each value represents the maximum power point determined every 5 min by a search algorithm for the maximum power point installed in an inverter.

The initial simulation parameters for a stage are set as follows: $i = 24$ such that for each i there is a 5 min interval where the AC–DC inverter algorithm determines the maximum power point data that is used to estimate the referential power curtailment $\rho_i^{**}(\delta_i)$. For a time interval of 5 s * 24 states, the analysis is made on what should be the referential working power for the maximum use of the weather conditions, setting a clarity index of kb $= 0.78$ for the next two hours ($120'$). In summary, the fundamental parameters used to obtain a referential power for each state are $\delta_1 = \theta_1$ (initial condition) where θ_1 is the first MPP captured by the investor, the number of states to use for prediction (in this case δ_{24}) such that for each i between 1 and 24 to obtain a referential power value that constitutes the working power $\rho_i^{**}(\delta_{i+1})$ in the next time interval.

For instance, the maximum power that the PVPP generates from 8:30 to 10:30 is plotted in Fig. 4. Applying the optimization tool created and considering the cloudiness, the new active power reference for these two hours is 0.5 p.u as it is illustrated in Fig. 4.

The following diagram in the algorithm in Table 2 shows in detail how the dynamic algorithm works to perform the power curtailment in such a way that once the overall optimal solution is known, any partial solution that involves only a part of the stages is also an optimal solution [16].

Considering this, the results for the complete day are illustrated in Fig. 5. The active power instead of being following the maximum power point follows a new

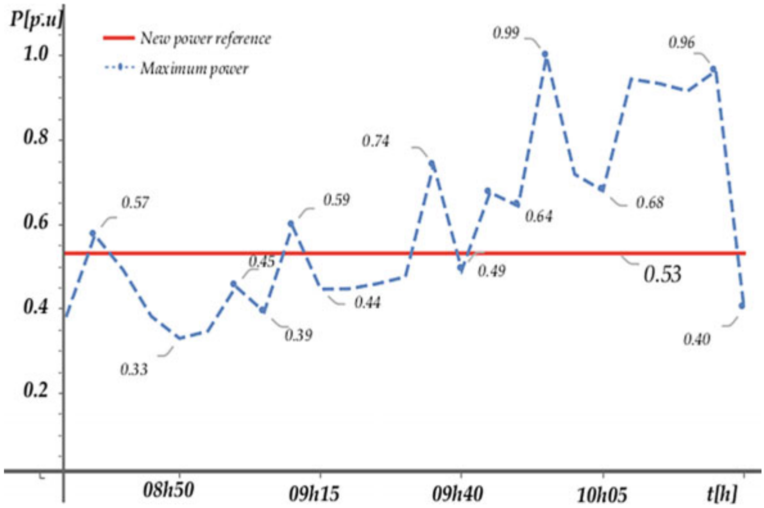


Fig. 4 Optimized active power from 8h30 to 10h30

Table 2 Algorithm for calculating the power curtailment

Algorithm	
Init in: Stage 1	• First time interval: $t = 5', i = 1$ • Initial decision: State $\delta_1 = \theta_1$ (Initial MPP value) • Save the historical power value by an accumulative mean: $C(\theta_1) = \theta_1$ • Find the optimal initial decision for next stage: $\boldsymbol{\rho_1^{**}(\delta_1, \theta_1) = Kb_1\left[C(\theta_1) + \rho_1^{*}(\delta_1, \theta_2)\right] = \delta_2}$
Stage 2	• Second time interval: $t = 10^{'}, i = 2$ • State 2 analysis: $\boldsymbol{\delta_2 = \rho_1^{**}(\delta_1, \theta_1)}$ (Optimal Decision in Last Stage) • Save the historical power value by an accumulative mean: $C(\theta_2) = \frac{(\theta_1+\theta_2)}{2}$ • Optimal decision for next stage: $\boldsymbol{\rho_2^{**}(\delta_2, \theta_2) = Kb_1\left[C(\theta_2) + \rho_1^{*}(\delta_1, \theta_3)\right] = \delta_3}$
⋮	⋮
End: Stage n	• General time interval: $\boldsymbol{t = 5' \cdot i, i = n}$ • State n analysis: $\boldsymbol{\delta_n = \rho_{n-1}^{**}(\delta_{n-1}, \theta_{n-1})}$ (Optimal Decision in Last Stage) • Save the historical power value by an accumulative mean: $C(\theta_n) = \frac{1}{n}\sum_{i=1}^{n}\theta_i$ • Optimal Decision for Last Stage: $\boldsymbol{\rho_n^{**}(\delta_n, \theta_n) = Kb_1\left[C(\theta_n) + \rho_1^{*}(\delta_n, \theta_{n+1})\right] = \delta_{n+1}}$

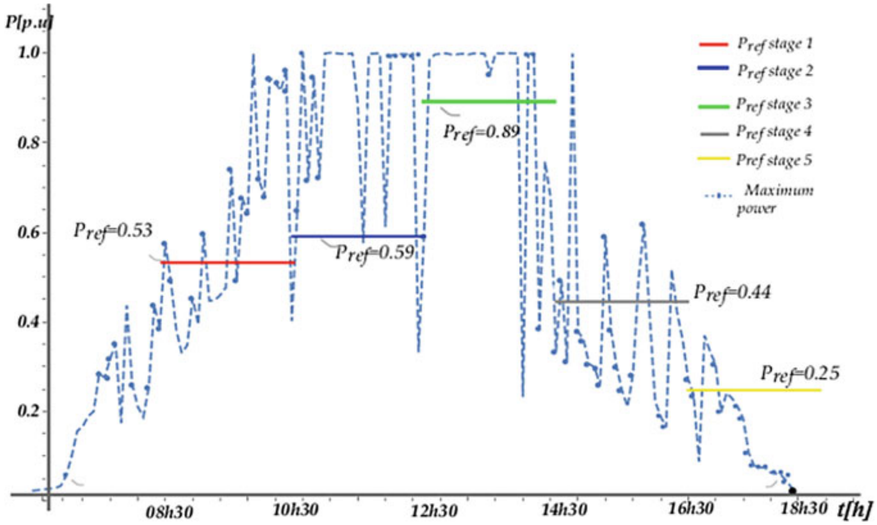


Fig. 5 Optimized active power generated for one day versus the maximum possible power

reference. This reference has been calculating by the model which considers the behavior on each part of the day and the cloudiness expected. As can be seen, when the PVPP uses this reference, the active power reduces its intermittent behaviors during the day.

The conditions that have been considered for decision making have been analyzed mainly considering that the state of the subsequent stages does not fully determine a decision policy of the current state, for this reason, it is considered of vital importance, the use of the clarity index (Kb) to conclusively determine the next state, which represents the optimal decision policy for each stage.

5 Conclusion and Future Scope

An optimization approach is to calculate the active power reference when power curtailment is employed. For this solution, stochastic programming has been handled. The optimization has been used considering one day data of a real photovoltaic power plant. The results determine the intermittent performance of the PVPP, which is reduced in each part of the day.

In the future scope, the optimization tool is considered, not only for solar irradiance and cloudiness but also for the economic aspects.

Acknowledgements The authors are greatly grateful for the support given by the SDAS Research Group. https://sdas-group.com/.

References

1. Cabrera-Tobar A, Bullich-Massagué E, Aragüés-Peñalba M, Gomis-Bellmunt O (2016) Review of advanced grid requirements for the integration of large scale photovoltaic power plants in the transmission system. Renew Sustain Energy Rev 62:971–987. https://doi.org/10.1016/j.rser.2016.05.044
2. Morais H, Kádár P, Faria P, Vale ZA, Khodr HM (2010) Optimal scheduling of a renewable micro-grid in an isolated load area using mixed-integer linear programming. Renew Energy 35(1):151–156. https://doi.org/10.1016/j.renene.2009.02.031
3. Fernández YF, Tobar AC, Peluffo-Ordóñez DH, Manosalvas TS, Miranda R (2019) Optimization-based algorithms applied in photovoltaic systems. RISTI Revista Iberica De Sistemas E Tecnologias De Informacao 2019(E22):242–255
4. Kaewpasuk S, Intiyot B, Jeenanunta C (2017) Stochastic unit commitment model for power system with renewable energy. In: 2017 International Electrical Engineering Congress (iEECON), pp 1–4 (2017). https://doi.org/10.1109/IEECON.2017.8075781
5. Papavasiliou A, Oren SS, O'Neill RP (2011) Reserve requirements for wind power integration: a scenario-based stochastic programming framework. IEEE Trans Power Syst 26(4):2197–2206. https://doi.org/10.1109/TPWRS.2011.2121095
6. Gevorgian V, Booth S (2013) Review of PREPA technical requirements for interconnecting wind and solar generation. Technical report, NREL, USA. https://doi.org/10.2172/1260328

7. Diniz AL, Maceira MEP (2013) Multi-lag benders decomposition for power generation planning with nonanticipativity constraints on the dispatch of LNG thermal plants. World Sci Ser Finance, Stochast Program 443–464. https://doi.org/10.1142/9789814407519_0016
8. Sasaki Y, Tsurumi T, Yorino N, Zoka Y, Rehiara AB (2019) Real-time dynamic economic load dispatch integrated with renewable energy curtailment. J Int Council Electr Eng 9(1):85–92. https://doi.org/10.1080/22348972.2019.1686861
9. Powell WB (2011) Approximate dynamic programming: solving the curses of dimensionality, 2nd edn. Wiley
10. Alvarez MAZ, Agbossou A, Cardenas A, Kelouwani S, Boulon L (2019) Demand response strategy applied to residential electric water heaters using dynamic programming and k-means clustering. IEEE Trans Sustain Energy 11(1):524–533. https://doi.org/10.1109/TSTE.2019.2897288.25
11. Haoxiang H (2019) Multi-objective optimization algorithm for power management in cognitive radio networks. J Ubiquit Comput Commun Technol 2:97–109. https://doi.org/10.36548/jucct.2019.2.004
12. Smys S, Raj JS (2019) Performance optimization of wireless Adhoc networks with authentication. J Ubiquit Comput Commun Technol 2:64–75. https://doi.org/10.36548/jucct.2019.2.001
13. Valsala KD, Premkumar K, Beevi AB (2019) Development of battery intervention power supply for solar roof top installations. Environ Progr Sustain Energy 38(2):570–583 (2019). https://doi.org/10.1002/ep.12958
14. Eltawil MA, Zhao Z (2013) MPPT techniques for photovoltaic applications. Renew Sustain Energy Rev 25:793–813. https://doi.org/10.1016/j.rser.2013.05.022
15. Dai Q, Fang X (2014) A simple model to predict solar radiation under clear sky conditions. Adv Space Res 53(8):1239–1245. https://doi.org/10.1016/j.asr.2014.01.025
16. Sniedovich M (1978) Dynamic programming and principles of optimality. J Math Anal Appl 65:586–606. https://doi.org/10.1016/0022-247X(78)90166-X

Arabic Braille Numeral Recognition Using Convolutional Neural Networks

Shurouq Alufaisan, Wafa Albur, Shaikha Alsedrah, and Ghazanfar Latif

Abstract Braille is a system that is designed to assist visually impaired individuals to acquire information. It consists of raised dots arranged in a cell of three rows and two columns. Visually impaired individuals rely on the sense of touch to read and write. However, it is difficult to memorize the arrangement of dots that compose a character. This research aims to design an application that recognizes and detects Arabic braille numerals and convert it to plain text and speech by implementing convolutional neural network variation Residual Network (ResNet). A new dataset was collected by capturing Arabic braille numerals using smartphone cameras. The recognition accuracy for Arabic braille numerals achieved 98%, taking into accountability different light and distance conditions.

Keywords Braille recognition · Deep learning · Arabic braille numerals classification · Convolutional neural network · Residual network

1 Introduction

Currently, everywhere in the world operate with the data as it is the most valuable part of our society, simply known as information. Collecting data successfully is the first step to operate effectively and make decisions efficiently. Sharing knowledge can be interpreted as communication, and it is one way of acquiring the needed information. In addition, reading is an important factor to learn and to obtain the information needed to prosper in our society. However, for visually impaired and deaf individuals, it is impossible to acquire information with plain texts, and there is no means of sharing information through communication with sighted individuals [1]. Therefore, a system, known as the braille system, was designed for visually impaired groups of individuals to access and receive information. With such a system, it is now possible to make decisions, operate upon events effectively, and communicate with sighted individuals.

S. Alufaisan · W. Albur · S. Alsedrah · G. Latif (✉)
College of Computer Engineering and Sciences, Prince Mohammad Bin Fahd University, Al Khobar, Saudi Arabia
e-mail: glatif@pmu.edu.sa

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_7

Braille system can be used as a mean of communication to share information. It is used for sighted individuals who wish to communicate with visually impaired individuals through written informational communication. Visually impaired individuals may prosper in our society and can play a significant role in our world. Thus, it is important to have a mean of communication between sight and visually impaired individuals to share information and to learn from each other.

Braille is a system that depends on the sense of touch. It is a system that enables visually impaired individuals to write and read with the help of the touch sense. It uses a series of raised dots that are used to read with the personal sense of touch. The language consists of six dots arranged in a rectangular shape. The dots can be arranged in any of the six position to compose a word, letter, number, or a special character. Braille system can be used to write different languages such as Arabic or English. In addition, the system can be used to write musical compositions and mathematical notations. Reading braille texts, for both Arabic and English, are read by moving the index finger from left to right.

Reading braille language with no previous experience can be difficult at first. To be able to read it effectively, the sense of touch must be trained first while making all the other senses unused. Also, users must memorize the positions of the dots and make sure what each composition of dots mean. Of course, not all individuals were born visually impaired. Some might go through events in life that make a person loses his/her sight, such as a chemical accident. In addition, there are situations, where a person might want to learn the language to teach it or to communicate with a visually impaired individual. Being a beginner in learning, this system can be hard at first. Ideally, if there exists a technology that can make reading and writing braille language easier, then it would be beneficial for the users to learn faster and to acquire the information that is needed efficiently. Such a system would have to translate the braille language into text and speech to be familiar with the system.

Therefore, the objective of this research is to develop an Arabic braille numeral system detection and recognition based on a deep learning algorithm. The proposed system will have the ability to recognize Arabic braille numeral images captured by cameras and to be processed by CNN variation, residual network [2, 3]. The goal of this work is to assist a person in reading and learning Arabic braille numeral. The ultimate goal is to have a pi camera that is integrated with the application, to scan and recognize printed Arabic braille numeral to translate it and display it to the user in plain text and speech [4–6]. This aims to improve the learning process of the Arabic braille system to acquire the required information for visually impaired users which lead to improve their daily life activities [7]. In addition, to help those who wish to communicate with visually impaired individuals.

2 Literature Review

Classifying and recognizing Arabic braille are significant to assist visually impaired individuals to learn and to obtain the necessary knowledge. In addition, to assist who wishes to communicate with a visually impaired individual. Extensive research has been conducted in the area of classifying and recognizing Arabic braille scripts. In [8], the authors proposed the use of find contour and artificial neural network for braille character recognition. Their method consists of preprocessing the image to prepare it for the process of the finding contour to get the black dots on the image for different datasets consisting of tilted images. The authors used segmentation to read the area of the braille cell. Later on, the artificial neural network was conducted as the final step to have the system learn by feeding it data input to obtain the desired data output value. The method achieved 99% of titled images of -1° to 1°. The authors showed that the level of accuracy decreased when the image is tilted more than 1°.

In [9], the author's dataset consisted of braille documents with the color of green and yellow that included dots in one side of the document. The author's method was to use image processing techniques to recognize Arabic braille characters. The authors preprocessing steps include: converting the image to grayscale, filtering the image, applying local thresholding for green braille documents, applying adaptive thresholding for yellow braille documents, segmentation, and extracting features. The authors described that Arabic braille characters were successfully recognized. Afterward, the authors proposed to convert Arabic braille character to binary strings which are converted into ASCII code to obtain the correct Arabic translation. The method achieved 98.04–100% for green braille document and 97.08–99.65% for yellow braille document. In [10], the authors proposed performing image preprocessing techniques to prepare the image for feature extraction. In the feature extraction step, the authors computed centroids of dots in the image to extract the relevant information from the image. Afterward, the authors aligned the coordination by applying many operations to rotate the centroids to align the page and the braille dots. Finally, the authors were able to recognize braille cells by grouping the dots to acquire a combination of letters and words. The method achieved between 94 and 99% braille cell recognition accuracy. In [11], the authors proposed converting the scanned document of braille page to gray color. After converting the image, the authors used the threshold technique to obtain three classed of regions. The authors used the three classes to initially identify braille dots. The possibility of valid dot identification was used in braille cell recognition. The method achieved 99% accuracy for skewed, reversed, or workout braille document.

In [12], the authors proposed classifying Arabic braille character using fuzzy classification, character feature extraction, and character recognition. The authors proposed system was developed with the use of segmenting Arabic braille characters. The authors use of fuzzy classification was inspired by the Fuzzy C-Mean (FCM) and fuzzy KNN classification algorithms. The method achieved up to 83% accuracy of classification and recognition of braille character. In [13], the authors suggest using a text-matching method as a way to recognize images. This system implies that starting with observing the interaction between words, then the use of several matching patterns between the phrases and ending up with matching entire sentences. This paper tried two methods to see what gives the highest accuracy, one of which was using paper citation matching, where the authors used a large academic dataset along with their citations and abstracts. The dataset sized 838,908 instances in total, containing 279,636 positive instance pairs and 559,272 negative instance pairs. The negative pairs were selected randomly, where they do not have citations along with them. Moreover, one out of three models that the authors trained outed the other two models. The three models being MP-IND, MP-COS, and MP-DOT, where MP-DOT was the model that gave the highest accuracy with a percentage of 88.73. In [14] seen that camera-enabled smartphone was used as the main method to capture braille characters. The paper suggested an algorithm that manipulates the images of braille documents that interpret the document's highlights and convert them to their equivalent English characters. Taking into account the lighting conditions while capturing the images, the authors obtained an accuracy of over 80% by developing an application that works under Android platform. The authors also stated that under the right conditions while capturing braille texts, the accuracy can go up to 84%.

In [15], the authors proposed a system for recognition double-sided Amharic braille documents that use the identification of three methods. Those methods are recto, verso, and overlapping dots. The system that the authors suggested in this paper works by simply converting braille texts into codes, and those codes are later on being translated into texts again. On top of that, adding the concept of reflection to reverse wrongly scanned braille documents automatically. While the dataset was collected from Addis Ababa University's Kennedy Library that contains good and bad scanned braille documents, the system was evaluated to give a high accuracy of 99.3% for identification and accuracy of 95.6% for translation. The authors in [16] used optical braille recognition based on semantic segmentation network along with auxiliary learning strategies. Using the OBR framework along with BraUNet and morphological post-processing procedures, the authors also used corresponding pixel-level annotations of braille characters with 64 other classes as an input in the system for both training and testing. The results of the methods they applied on DSBI dataset, type recto braille character along with BraUNet gave the best results with an accuracy of 0.9966%. While the regular number of classes for braille classes is 64 the authors in [17] took another approach and added 7 more classes to add up to 71 classes of characters to corresponding to the braille dataset that consists of 37 characters. Moreover, a collective dataset of 26,724 labeled braille images now has 37 braille symbols that correspond to the 71 classes. Using a novel method that pairs ratio character segmentation, RCSA's algorithm was used aside with CNN to translate a line of braille into its English counterpart. By linking the CNN model to two recognition techniques: character and word recognition the system proposed in this paper were able to reach an accuracy of 98.73% on the test set. In [18], the authors used Convolutional Neural Network (CNN) techniques to develop a system that can identify Cyrillic braille characters. After scanning the braille documents, image preprocessing techniques were used to make the recognition process easier. Then, character segmentation was performed to improve recognition accuracy. Subsequently, a modified backpropagation algorithm was used to train neural networks. The system has achieved 95.7% training accuracy and 95% testing accuracy. The authors also concluded that the use of the artificial neural network is very helpful in identifying characters due to the ease of programming the network architecture to train and test with any image sizes as an input. In [19], the authors proposed a module that uses associative memories to recognize single-sided braille documents and then convert it to audio. Their module consisted of two stages, preprocessing and recognition. In the preprocessing stage, different operations are performed on the scanned braille papers such as grayscale conversion and dilation to prepare the images for the next stage. Then, Modify Multi-Connect Architecture (MMCA) and Modify Bidirectional Associative Memory (MBAM) algorithms are used to recognize the characters. The authors compared results of MMCA and MBAM algorithms, where (MMCA) achieved an average accuracy of 98.26% for characters recognition while (MBAM) achieved 91.87%. Afterward, the proposed module converts the recognized text into audio.

In [20], the authors proposed developed an Optical Braille Translator (OBT) system that identifies single-sided Sinhala braille characters and translates it into texts. The systems features were developed based on image processing techniques in MATLAB environment. Their methodology was to apply preprocessing techniques like grayscale conversion, image rescaling, and angle correction functions on both handwritten and computer-generated braille documents. Then, segmentations are used to recognize the braille character cells. Afterward, extracted characters are resized into a 21 × 16 matrix binary images. Braille characters are regenerated using an algorithm to improve accuracy. The developed system was able to reach an accuracy of over 99%. In [21], the authors proposed a system equipped with a scanner and a webcam. The braille characters are captured by the webcam or the scanner. Then, grayscale conversion, filtering, segmentation, and normalization are done on the captured images using MATLAB IDE. After the image processing stage, Artificial Neural Networks (ANN) and feature extraction are used to recognize the patterns of braille characters as well as obtaining a training model as a dataset. Furthermore, the recognized characters are then converted to audio. The system results showed that the characters captured with webcam resulted in 57.69% accuracy while the characters captured using the scanner resulted in an average of 93.26% accuracy. In [14], the authors proposed a method to process braille documents and convert them into English text and speech. Their method performs Hough's circle detection algorithm on the phone captured braille images to identify the dots. Then, maximum length and width are calculated for all dots to generate a new image of equaled size dots. After that, the authors used image segmentation to divide the rows and columns into cells. Later on, each cell in a row is read to recognize its pattern by the position of its dots. Finally, each cell pattern is mapped to its matched English character. The output text is converted into speech using a text-to-speech engine. In ideal lighting and alignment conditions, the method achieved more than 80% accuracy.

3 Methodology

The overall aim of this research is to design an application that uses deep learning techniques that recognize Arabic braille numerals with high accuracy. Captured images of Arabic braille numerals are used to train the CNN variation Residual Network (ResNet). The model was modified to achieve high accuracy of recognizing Arabic braille numerals.

Training a deep network is a challenge, and it has been proven that the depth of a network degrades the network performance [22, 23]. To address this problem, ResNet is chosen as the building block of our network due to its methods of training a deep network efficiently. Adding more layers to the network leads to the vanishing gradient issue which shows high training error. Thus, the authors in [24] suggested adding skip connections that skip one or more layers. The authors in [24] proved that when the model uses skip connection, the training efficiency improved since the gradients can travel to many layers with the help of skip connections. The proposed model is implemented using python libraries that will accept images and classify them to its respective classes. The application will be able to detect and recognize Arabic braille numerals using a pi camera that will be integrated with raspberry pi 4. Once the model is trained, captured images of Arabic braille numerals will be classified to its respective classes, provide a correct translation of it and then convert it to speech audio representing the numeral that was processed.

3.1 Data Description

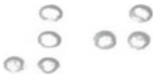
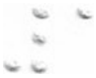
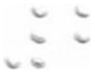
Arabic braille numerals were printed on a single side A4 embossed paper with color blue and white documents. Braille dots are arranged in a cell with three rows and two columns. Dot height is approximately 0.02 inches, and the horizontal and vertical spacing between the dots are 0.1 inches. The blank space between the cells is 1.3 inches.

As Table 1 shows, 5000 images of Arabic braille numerals were collected. Smartphone cameras are used to generate the dataset. The images are captured under different lights, such as natural and industrial lighting. The images range from several colors, white, yellow, and gray. Images were captured on all possible angles and heights.

3.2 Preprocessing

In this stage, images are being prepared to make it easier for the model to train and to recognize braille numerals.

Table 1 Arabic braille numerals sample dataset

#	Name in Arabic	Name in English	Sample image	Number of images
0	صفر	Sifer		500
1	واحد	Wahed		500
2	اثنين	Ethnein		500

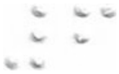
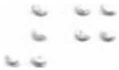
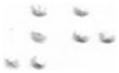
(continued)

Table 1 (continued)

#	Name in Arabic	Name in English	Sample image	Number of images
3	ثلاثة	Thalatha		500
4	اربعة	Arba-a		500
5	خمسة	Khamsa		500

(continued)

Table 1 (continued)

#	Name in Arabic	Name in English	Sample image	Number of images
6	ستة	Sitta		500
7	سبعة	Sab-a		500
8	ثمانية	Thamanya		500

(continued)

Table 1 (continued)

#	Name in Arabic	Name in English	Sample image	Number of images
9	تسعة	Tis-a		500

3.2.1 Grayscaling

Computers identify RGB images as 3D arrays. Meanwhile, grayscale images are identified as 2D arrays, which make the preprocessing stage more efficient. Thus, captured RGB images of braille numerals are converted to grayscale.

3.2.2 Resizing

Training images of larger size will add computational power to the network as well as time complexity. Therefore, images are resized to 256, along the y and x-axis, before inputting them into ResNet. With this approach, the model will be able to train faster with less pixels, while preserving important features.

3.2.3 Converting to Array

In deep learning, the model is trained with images in a NumPy array form. Thus, grayscaled images will be converted to an array using the NumPy library. Later on, the arrays will be fed to the network to start the training.

3.3 Deep Residual Learning

The depth of a network is significant in any neural network model [25]. However, the deeper the network the harder it is to train the model. It is noticed that as the network gets deeper, there will be a degradation problem which will lead to a decrease in the validation set. In residual block, the layers would be fed into the next layer and the layer after it, while having the advantage to skip the number of layers known as skip connection. The diagram below (Fig. 1) illustrates the skip connection.

The diagram on the right describes a deep network that shows stacked layers one after the other. However, the diagram on the left describes a deep network with stacked layers as before but the original input is now added to the output of the block. This operation is called a skip connection. Skipping the training of a few layers can be beneficial in solving the degradation problem that affects the accuracy negatively due to having more layers than needed for training.

In [17], the authors suggested using the building block, which is defined as:

$$y = F(x, \{W_i\} + x \tag{1}$$

The above equation is the shortcut connection which does not hold extra parameters and computation complexity. The identity mapping is responsible for adding the output from the previous layer to the next layer. Input and output vectors are represented as x and y of the considered layers. $F(x, \{W_i\}$ is the function that is used

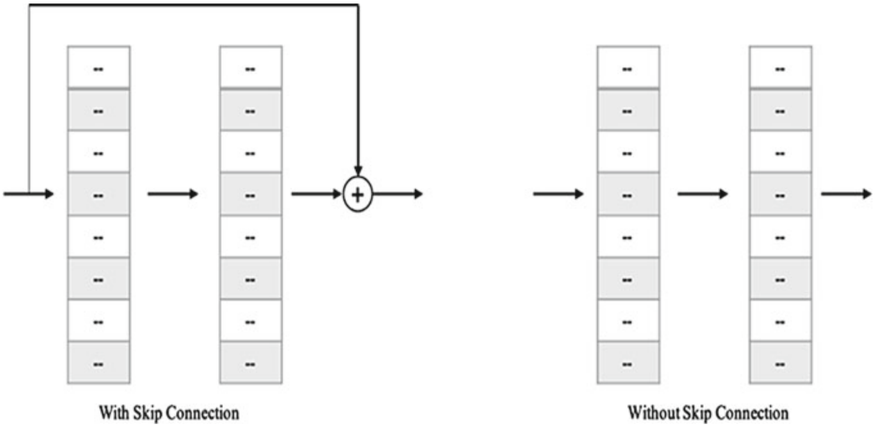


Fig. 1 Skip connection concept

for the residual mapping that will be learned. It is a requirement for the input and output to have equal dimensions to use the shortcut connection equation. Also, in [17], the authors discussed a solution in case when the dimensions are not equal that is linear projection. If the model found that the dimensions are not equal, then it will have an operation which will multiply a linear projection W to the identity mapping to match the dimensions. The equation is defined as:

$$y = F(x, \{W_i\} + W_s x \tag{2}$$

while identity mapping is used to solve the degradation problem multiplying is W_s used when there is a need to match the dimensions [17].

4 Results and Discussion

In this section, the performance of CNN was examined for training and recognizing Arabic braille numerals. A modified ResNet architecture model was used to implement our system. The dataset has been divided based on their corresponding labels to guarantee that each image will be classified based on their labels. The model was trained for 120 epochs on the braille numerals dataset, which means training for 10 classes for 120 cycles. 20% of the dataset images were used for validation and testing. This means that 3000 images were used for training, 1000 for validation, and 1000 for testing. The results using the ResNet model were extremely high, the model achieved a validation accuracy of 98%, as long with high model performance. Figure 2 shows the validation accuracy and test accuracy.

Table 2 shows a comparison between the results of the first and last epoch. The validation has increased from 0.40 to 0.98 as the training has progressed.

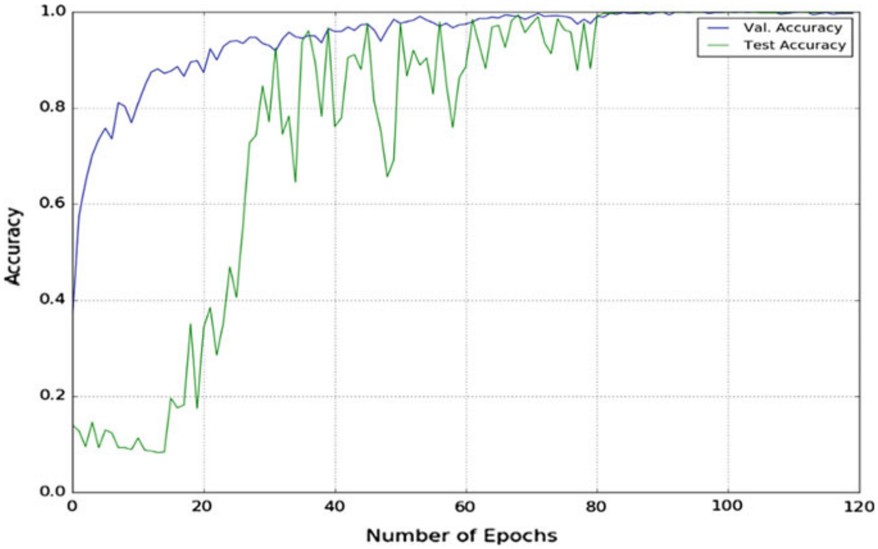


Fig. 2 Model validation and testing accuracy curve

Table 2 First and last epoch results

Epoch no.	Accuracy	
	Test	Val
First epoch	0.14	0.40
Last epoch	0.99	0.98

From Fig. 2 can be observed that the test accuracy has dramatically increased after epoch no. 25.

5 Conclusion

In this article, Arabic braille recognition using deep learning techniques was proposed. Braille is a system that is designed to assist visually impaired individuals to acquire information. Visually impaired individuals rely on the sense of touch to read and write. The proposed system can help teachers, parents, and people who have lost their vision recently. The system uses a modified ResNet model, which overcomes the degradation problem. A validation accuracy of 98% and a test accuracy of 99% were able to achieve. Deep learning techniques have proved their efficiency in recognizing braille characters.

References

1. Latif G, Mohammad N, AlKhalaf R, AlKhalaf R, Alghazo J, Khan M (2020) An automatic arabic sign language recognition system based on deep CNN: an assistive system for the deaf and hard of hearing. Int J Comput Dig Syst 9(4):715–724
2. Alghmgham DA, Latif G, Alghazo J, Alzubaidi L (2019) Autonomous traffic sign (ATSR) detection and recognition using deep CNN. Procedia Comput Sci 163:266–274
3. Shaikh E, Mohiuddin I, Manzoor A, Latif G, Mohammad N (2019) Automated grading for handwritten answer sheets using convolutional neural networks. In: 2019 2nd International conference on new trends in computing sciences (ICTCS), pp 1–6. IEEE, Oct 2019
4. Mahmoud AA, Alawadh INA, Latif G, Alghazo J (2020) Smart nursery for smart cities: ınfant sound classification based on novel features and support vector classifier. In: 2020 7th ınternational conference on electrical and electronics engineering (ICEEE), pp 47–52. IEEE, Apr 2020
5. Al-Hmouz A, Latif G, Alghazo J, Al-Hmouz R (2020) Enhanced numeral recognition for handwritten multi-language numerals using fuzzy set-based decision mechanism. Int J Mach Learn Comput 10(1)
6. Alzubaidi L, Latif G, Alghazo JM, Zikria M (2019) Cloud-based ınteractive hands free e-learning environment for students with disabilities
7. Latif G, Alghazo JM, Maheswar R, Jayarajan P, Sampathkumar A (2020) Impact of IoT-based smart cities on human daily life. In: Integration of WSN and IoT for smart cities. Springer, Cham, pp 103–114
8. Subur J, Sardjono TA, Mardiyanto R (2016) Braille character recognition using find contour and artificial neural network. J Electr Electron Eng 14(1)
9. Authman ZI, Jebr ZF (2012) Arabic Braille scripts recognition and translation using image processing techniques. J Educ Pure Sci 2(3):18–26
10. Mousa A, Hiary H, Alomari R, Alnemer L (2013) Download limit exceeded, citeseerx.ist.psu.edu, Nov 2013. Available: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.403.6856&rep=rep1&type=pdf.
11. Al-Salman A, AlOhali Y, AlKanhal M, AlRajih A (2007) An Arabic optical braille recognition system
12. Al Nassiri A, Abdulla S (2018) A fuzzy classification and recognition system for Arabic Braille segmented characters. Int J Appl Eng Res 13(6):3662–3669
13. Pang L, Lan Y, Guo J, Xu J, Wan S, Cheng X (2016) Text matching as image recognition. arXiv preprint arXiv:1602.06359
14. Venugopal-Wairagade GA (2016) Braille recognition using a camera-enabled smartphone. Int J Eng Manuf 6(4):32–39
15. Seid H, Assabie Y (2017) Recognition of double sided Amharic Braille documents recognition of double sided Amharic Braille documents view project NLP tools for Ethiopian languages view project recognition of double sided Amharic Braille documents. Int J Image Graph Sig Process 4:1–9
16. Li R, Liu H, Wang X, Xu J, Qian Y (2020) Optical Braille recognition based on semantic segmentation network with auxiliary learning strategy. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp 554–555
17. Hsu B-M (2020) Braille Recognition for reducing asymmetric communication between the blind and non-blind. Symmetry 12(7):1069
18. Smelyakov K, Chupryna A, Yeremenko D, Sakhon A, Polezhai V (2018) Braille character recognition based on neural networks. In: 2018 IEEE second ınternational conference on data stream mining & processing (DSMP), pp 509–513. IEEE, Aug 2018
19. Khaled S, Safana H, Abbas (2017) Braille character recognition using associative memories. Int J Eng Res Adv Technol (IJERAT) 1:2454–6135
20. Perera TDSH, Wanniarachchi WKIL (2018) Optical Braille translator for Sinhala Braille system: paper communication tool between vision ımpaired and sighted persons. Int J Multimed Its Appl (IJMA) 10

21. Ramiati, Aulia S, Lifwarda, Nindya Satriani SN (2020) Recognition of ımage pattern to ıdentification of braille characters to be audio signals for blind communication tools. İn: IOP conference series: materials science and engineering, vol 846, p 012008, May 2020
22. Chen Z, Xie Z, Zhang W, Xu X (2017) ResNet and model fusion for automatic spoofing detection. In: INTERSPEECH, Aug 2017, pp 102–106
23. Latif G, Iskandar DNFA, Alghazo J, Butt MM (2020) Brain MR ımage classification for glioma tumor detection using deep convolutional neural network features. Curr Med Imag
24. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition, 2016, pp 770–778
25. Latif G, Alghazo J, Maheswar R, Vijayakumar V, Butt M (2020) Deep learning based intelligence cognitive vision drone for automatic plant diseases identification and spraying. J Intell Fuzzy Syst (Preprint) 1–12

In-Silico Modeling of Sleep Stage Classification System Using Covariance Normalization

Nishant Gautam, G. B. Mrudula, and C. Santhosh Kumar

Abstract With surge among count of sleeping disorders across the globe and among every strata of society and the non-availability of sleep medicine tools in the backward regions of the third world nations, the need of automated systems arises. This paper introduces an inexpensive, computerized binary sleep stage classification system classifying the rapid eye movement (REM) and non-rapid eye movement (NREM) stages with the usage of electrocardiogram (ECG) and respiratory effort signals (RES). To avail a baseline classification of the sleep stages, support vector machine (SVM) was employed as the backend classifier that uses the heart rate variability (HRV) and respiratory rate variability (RRV) features derived from ECG and RES, respectively. The baseline system developed using linear SVM kernal performed better with performance accuracy of 73.83% , sensitivity of 84.37% and specificity of 30% in totality. The statistical features extracted from the data contain patient-specific variations that are irrelevant for sleep stage classification. As an effort to minimize these variations, covariance normalization (CVN) was performed, and a system is obtained with an absolute overall classification accuracy of 81.30%.

Keywords Electrocardiogram · Respiratory effort signals · Respiratory rate variability · Heart rate variability · Covariance normalization

---

N. Gautam (✉) · G. B. Mrudula · C. Santhosh Kumar
Machine Intelligence Research Laboratory, Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: g.nishant12329@gmail.com

G. B. Mrudula
e-mail: gb_mrudula@cb.students.amrita.edu

C. Santhosh Kumar
e-mail: cs_kumar@cb.amrita.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_8

1 Introduction

Sleep [1], the ingrained natural, easily reversible periodic state of all the living beings marked by the state in between wakefulness and loss of consciousness of one's surrounding, is accompanied by the typical body posture (lying down, closed eyes), occurrence of dreams, changes in brain's electrical activity, reduced muscle activity and other physiological signaling. Sleep comprises recurrent stages of NREM and REM within the sleep cycle and is considered to be highly important for the restoration and recovery of vital bodily and cerebral functioning. The electrophysiological variables associated to the sleep study are electroencephalography (EEG), electrocardiogram (ECG), electrooculogram (EOG), chin electromyogram (Chin-EMG), respitory effort signals (RES) and oxygen saturation (SpO_2) [2]. Sleep is greatly affected by the circadian cycle managed by cerebral neurons, which collectively builds up the sleep cycle of human body.

Traditionally, sleep cycle is sorted out into two stages, non-rapid eye movement (NREM) and rapid eye movement (REM). Both of these differentiated using not just neuro-physiological but also psycho-physiological characteristics. NREM and REM stages of sleep alternate in a cyclic manner, with a total of 4–6 cycles, each cycle lasting for 90–110 min. The NREM sleep often referred as quiet sleep, which continues as REM sleep also known as active sleep is subdivided to foour stages. Initial stage of NREM sleep is described as a transition time from wakefulness to sleep. As NREM sleep advances, the heartbeat gradually slows down, muscle activity decreases, the consciousness of the individual fades out completely, and the sleep cycle enters into the REM stage. The REM stage is a sleep period with intense brain activity, where muscles are inactive and unresponsive characterized by rapid movements of eye in various directions. Sleep disorders like sleep apnea are very common to occur in this stage due to the lose of muscle tone. During REM sleep, there is an overall increase in breathing rate, and there is a more pronounced variation in cardiovascular activity, with overall increases in blood pressure and heart rate and also ST segment elevation. Additionally, changes in blood flow that cause erections to occur in males or swelling of the clitoris in females are characteristic of REM sleep. The underlying reason for these considerable cardiac and physiological variations in REM sleep is currently unknown and may be a by-product of REM-related changes in nervous system activity or related to dream content. In the contemporary world, large population suffers from the sleeping disorders [3] irrespective of age and gender, like sleep apnea, insomnia, restless leg syndrome (RLS) and others. Limitations of the traditional non-invasive sleep diagnostic tool polysomnography (PSG) [4] are that, to avail the electrophysiological data, a complex array of electrical sensors is attached to the body making patient uncomfortable. Also, the procedure is time consuming and requires an expert neurotechnician to interpret the results and proffer the diagnosis which makes the sleep medicine an expensive and arduous task. Hence, the non-familiar environment and the possibility of interpretation error makes the data erroneous that lead to having a great potential of misdiagnosis. Therefore, the need of development of an automated, portable, cheap sleep state classification system arises.

The foundation of this cardiorespiratory system resides on H.R Variability and R.R Variability features [5, 6]. Till date several researches have been carried out in the sleep stage classification domain in which the features were derived from varied sleep stages and substages within them. Hence, our research work eliminates the particular stage and the varied dependencies of feature values. HRV, aka cardiac beat by beat variation and is highly necessary to identify the autonomic vitality defined by the parasympathetic superiority within NREM stage also with a shift to sympathetic ascendency within REM stage. RRV is also an important criterion for different humane parameters and system. RRV varies greatly with the sleep stages matching up with the defined depths of sleep and posseses greater utility when a person is awake. Hence, the proposed system employs only ECG and RES signals, making the system effective and accessible [7] to use in backward localities having only primary healthcare furnished with the simple data acquisition systems [8]. Sleep impoverishment is associated with alteration in HRV and RRV, which makes these features a promising identifier for sleep disorders and sleep stages.

2 Methodology

2.1 Dataset

PSG signals were acquired from the Department of Neurology at Amrita Institute of Medical Science (AIMS), Kochi. The dataset comprised of full single night PSG recordings of total 32 subjects aging from 17 to 75 years. The sleep study data comprised of physiological parameters like both side occipital EEG, SpO_2, ECG, RES, nasal intake along with left and right side EOG. The final dataset constructed comprised of 247 samples, out of which the division of data for testing and training was achieved. The data was split into 70:30 ratio, hence 173 samples were kept for training, and 74 samples were maintained for testing purpose.

2.2 Feature Extraction and Feature Fusion

Signals were processed in advance, earlier than extracting needed features for the study. A 50 Hz notch filter was used to filter out the power line interference from ECG. Artifacts from effort signals were removed using a filter having a cut-off frequency of 1 Hz. R-R [9] and breath to breath interval was drawn out from effort signals and cardiogram signal. HRV and RRV features in both time domain and frequency domain were extracted from the interval data. Feature extracted was fused together through feature fusion [10]. Hence, a combined featured data comprising fused time and frequency domain features was also constructed for the study.

- **R-R Interval**
  Time elapsed time between two successive R waves within the QRS complex on the electrocardiogram has unit of measurement in seconds. General (normal) value of RR interval is 0.6–1.2 s. And any sort of variability besides the mentioned normal value reflects the disorder within the subject.
- **B-B Interval**
  Breath to breath interval [11] was monitored within the breath cycle of the test subjects to analyze the patient's degree of breathing while she/he is asleep. All breath to breath intervals were automatically analyzed from flow signal, displayed and manually corrected for artifacts. Variations among the B-B interval during sleep cycle can indicate toward the morbidity within that patient.

Time Domain Features (TD) Refers to variation of amplitude of signal with time. Here in this domain features extracted encompass mean, standard deviation and root mean square successive difference (RMSSD) of R-R and B-B intervals, mean heart and breathing rates, percentage of the mean number of times an hour in which the change in successive normal sinus (NN) intervals exceeds 50 ms. Above them all, the data from respiratory inductance plethysmography (RIP) was involved for evaluating pulmonary ventilation by taking the movement of chest wall and abdomen into the equation. This feature known as rib cage percentage (RC%) is defined as the contribution (in percentage) of the rib cage to the inspiratory tidal volume, given by the equation:

$$\mathrm{RC\%} = \frac{\text{Thoracic Movement}}{\text{Thoracic Movement} + \text{Abdomen Movement}} \tag{1}$$

Frequency Domain Features (FD) The frequency domain refers to the analytic space in which mathematical functions or signals are conveyed in terms of frequency, rather than time. Features of both HRV and RRV were considered for this respective domain, and therefore, three frequency bands given in Tables 1 and 2 were considered and employed. Frequency domain features encompass absolute power, maximum frequency, peak power, ratio among low frequency and high frequency.

Table 1 HRV frequency bands

Band name	Abbreviation	Wavelength (Hz)
Very low frequency	VLF	0.0003–0.04
Low frequency	LF	0.04–0.15
High frequency	HF	0.15–0.4

Table 2 RRV frequency bands

Band name	Abbreviation	Wavelength (Hz)
Very low frequency	VLF	0.0003–0.05
Low frequency	LF	0.05–0.2
High frequency	HF	0.2–0.25

2.3 Baseline System

Baseline system which is to act as reference point of this study was created using the features extracted from the ECG and RES data of all subjects. The database then later on separated into the training and testing set, and the SVM [12]-based backend classifiers were used for the training the classifier and to develop the SVM-based sleep stage classification model. Representation of the baseline/reference system is given in Fig. 1.

Polysomnography (PSG) Polysomnography, a diagnostic tool to analyze patients having a potential sleeping disorder, records the multiple parameters that are associated with the sleeping disorders such as heart rate, EEG, skeletal muscle activity (EMG), oxygen saturation, eye movement (EOG) and also the respiratory effort indicators. Test is performed within sleep lab in the presence of an expert technologist, who gives actual diagnosis after a step of sleep stage scoring which is performed by epoch method through assessment of sleep parameters.

Here, in our research work instead of using all the extracted parameters, only ECG and RES signals were employed and processed.

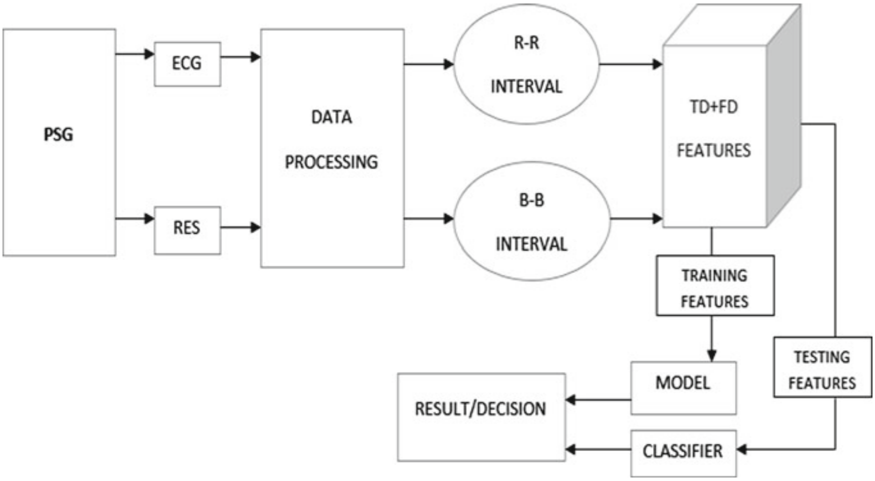


Fig. 1 Baseline system

Electrocardiogram (ECG) An electrocardiogram records the heart's electrophysiological data abbreviated as EKG and ECG. ECG shows graph of voltage vs time of cardiac bio-potential. Traditionally, it is of 12 leads out of which ten were placed on patient's limbs and chest. ECG is made up of frequency peaks, intervals and segments named PR interval, PR segment, QRS complex, ST segment, etc., each having a unique significance and importance when vitality of the heart is in question.

Respiratory Effort Signal (RES) RES was measured using sensor bands placed around thoracic and abdominal regions, and respiratory airflow was measured using thermocouple sensors placed at the nasal and buccal offices. It involves the procedure of respiratory inductance plethysmography which measures the pulmonary ventilation by evaluating chest and abdominal movements. Although number of complex respiratory parameters such as tidal volume (vt), minute ventilation, respiratory rate can be obtained, but here for our research purpose, RC% (rib cage percentage ratio) was considered.

Data Processing Raw data was processed by employing different filters (LPF and notch filters). Artifacts were removed from the ECG and RES signals, and the time domain and frequency domain features were extracted (refer Sect. 2.2). Post data processing features extraction procedure was employed that gives rise to our final dataset which further was divided into testing and training data (refer Sect. 2.1).

Model and Classifier SVM backend classifier was employed over the final dataset, and implementation of four different kernels (linear, polynomial, rbf and sigmoid) was performed over the datasets (test and training). Finally, best performance was noted down and taken as the result.

2.4 Covariance Normalization

A feature normalization [13] approach safeguards the inherent correspondence among the feature vectors. In real-time data acquisition, there is a presence of specific attenuation to the physiological signals of each subject. This normalization contains the ability to minimize the noisy factor established due to subject-specific variations. Stepwise procedure of covariance normalization (CVN) [14] is given in Fig. 2.

Covariance Matrix Covariance matrix is a square matrix providing the covariance between each pair of elements of a given feature vector. Covariance gives the joint variability of two random variables. It measures the direct relationship between two quantities.

$$\mathrm{Cov}(X, Y) = \frac{\sum(Xi - \bar{X})(Yi - \bar{Y})}{N} \tag{2}$$

In true physical sense, if covariance is positive, the quantities related move together in a proportional fashion, on the other hand, if the covariance is negative, the quantities are inversely proportional to each other. Covariance matrix is positive semidefinite

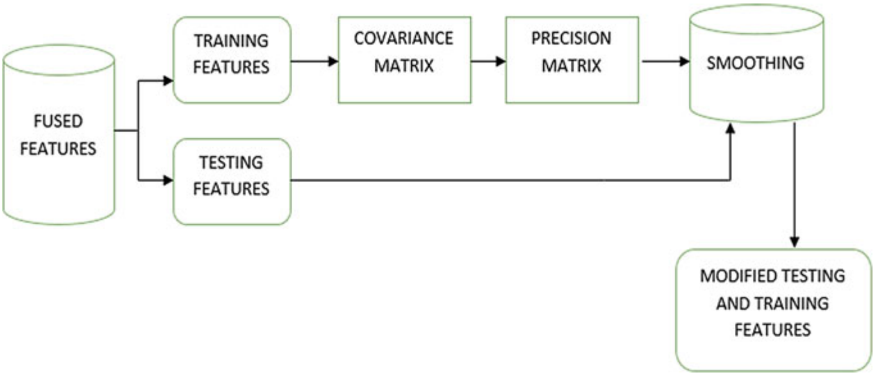


Fig. 2 CVN block diagram

matrix having the eigenvalue positive and the scalar strictly positive for every non-zero column vector z of n real numbers. Here, the covariance was employed over the training data resulting in the matrix formulation.

Precision Matrix Commonly known as inverse covariance matrix or concentration matrix. In true physical sense, it is identified as the converse of covariance matrix. Reason to obtain the concentration matrix is its association with the multivariate normal distribution and partial correlation.

Data Smoothing Estimates obtained after covariance are considered as noisy, which are smoothened out by computing smoothening matrix. This allows important patterns to stand out while leaving noise out. Smoothing can be used to help predict trends. Here, data is compiled, and it can be manipulated to remove or reduce any fallacy or noise.

Here, the smoothing of the data is performed using the equation:

$$S = \lambda P + (1 - \lambda)I \tag{3}$$

within the equation S is smoothing matrix, P is upper triangular matrix extracted from cholesky factorization, lambda is the smoothing factor, and identity matrix is represented with a symbol I. The main motive to smooth the data is to identify simplified alterations in order to help predict different trends and patterns. The smoothened training features were obtained and were fed to the classifier to develop the CVN-based sleep stage classification system.

3 Results and Discussion

Here, in the results, the best performance out of the four kernels employed was taken into the account.

3.1 Support Vector Machines (SVMs)

Support vector machine is a widely used supervised machine learning model useful for classification approaches and works effectively for two stage classification problems. After feeding an SVM model sets of labeled training data for each category, they are able to categorize or classify the dataset efficiently into binary classes. A support vector machine takes the data values and outputs the hyperplane (which in two dimensions is simply a line) that best separates the labeled data. Hyperplane is a hypothetical decision boundary that separates the data points into the feature space: Anything that falls to one side will classify as the class 0 and other will be class 1. LIBSVM [12] package was employed over the Python platform and thereby usage of four kernels (shown below). And the accuracy of each one was noted.

- **Linear Kernel**.
  Simplest of them all, this function is the inner product of x&y added with a constant value (c). This function separates the data into two classes using a linear function given by:
  $$P(x, y) = (x^T * y) + c \tag{4}$$
- **Polynomial Kernel.**
  It is a non-stationary function, suitable for the situations where the data is pre-normalized. Represented as:
  $$P(x, y) = (ax^T * y) + c \tag{5}$$
- **RBF (Radial Basis Function Kernel).**
  It is nonlinear function that separates the data into classes, used when the boundaries are hypothesized and have a curvy fashion.
- **Sigmoid Kernel.** Also known as hyperbolic tangent kernel. This function placed itself under the neural networks learning field, where it is used as an activation function for the perceptrons.
  $$P(x, y) = \tan h(ax^T * y + c) \tag{6}$$

Baseline System System's performance for TD, FD and TD-FD features is tabulated in Table 3. The baseline performance was seen best for linear kernel-SVM.

Table 3 Baseline system performance (%)

Input features	Accuracy	Sensitivity	Specificity
TD	73.83	78.30	30
FD	63.5	84.37	32.55
TD and FD	57	77.61	22.5

Table 4 CVN system performance (%)

Input features	Accuracy	Sensitivity	Specificity
TD	80.373	81.63	66.6
FD	81.30	81.80	75
TD and FD	81.30	80.50	100

Covariance Normalization System The performance accuracy of the normalized system for all the features is tabulated in Table 4. The performance of the system was seen best for linear kernel-SVM.

4 Conclusion

Within this work, an approach of covariance normalization was explored to reduce the disparites associated to a particular subject, embedded within features in order to ameliorate the execution of the sleep stage classification system. Reference system was created using an SVM classifier. The statistical features extracted from the heart's electrophysiology and effort signals were fed to the baseline/reference system as an input. Reference system showed a performance of 73.83% for time domain features, 63.5% for frequency domain features and 57.0% for both time domain and frequency domain features combined. Further, the patient- and stage-specific variations of the features were reduced using a feature normalization technique called covariance normalization. The performance of the CVN-SVM system with TD features as input had a performance improvement of 6.54% absolute, CVN-SVM system with FD features as input, had performance improvement of 17.8% absolute and CVN-SVM system with both TD and FD features as input had a performance improvement of 24.3% absolute. This cardiorespiratory system could be helpful to the patients suffering from sleeping disorders like sleep apnea, restless leg syndrome and others. CVN approach favored the aim and needs of the study and performed better than the baseline system, with the best performance of 81.30%

Acknowledgements My deepest and unfeigned appreciation to my friends and faculties for the support and guidance throughout the research work. Also, I sincerely express my humble gratitude to Haritha H and Sreekumar K T of Machine Intelligence Research Laboratory (MIRL) for providing essential technical support and the needed motivation all along.

References

1. Borb A, Achermann P (1999) Sleep homeostasis and models of sleep regulation. J Biol Rhyth 14(6):559–570
2. Redline S et al (1998) Methods for obtaining and analyzing unattended polysomnography data for a multicenter study. Sleep 21(7):759–767

3. Sateia M (2014) International classification of sleep disorders-third edition. Chest 146(5):1387–1394
4. Douglas N, Thomas S, Jan M (1992) Clinical value of polysomnography. Lancet 339(8789):347–350
5. Stein PK, Pu Y (2012) Heart rate variability, sleep and sleep disorders. Sleep Med Rev 16(1):47–66
6. Gutierrez G, Williams J, Alrehaili GA, McLean A, Pirouz R, Amdur R, Jain V, Ahari J, Bawa A, Kimbro S (2016) Respiratory rate variability in sleeping adults without obstructive sleep apnea. Physiol Rep 4(17)
7. Joby PP (2019) Exploring devops: challenges and benefits. J. Inf. Technol. Digital World 01(01):27–37
8. Shakya D (2020) Analysis of artificial intelligence based image classification techniques. J Innov Image Process 2(1):44–54
9. Manikandan MS, Soman K (2012) A novel method for detecting r-peaks in electrocardiogram (ECG) signal. Biomed Signal Process Control 7(2):118–128
10. Yang J, Yang J, Zhang D, Lu J (2003) Feature fusion: parallel strategy versus serial strategy. Pattern Recogn 36(6):1369–1381
11. Kowallik P, Jacobi I, Jirmann A, Meesmann M, Schmidt M, Wirtz H (2001) Breath-to-breath variability correlates with apnea-hypopnea index in obstructive sleep apnea. Chest 119(2):451–459
12. Chang C-C, Lin C-J (2011) Libsvm. ACM Trans Intell Syst Technol 2(3):1–27
13. Kumar CS, Ramachandran KI, Kumar AA. Vital sign normalisation for improving performance of multi-parameter patient monitors. Electron Lett 51(25)
14. Mrudula GB, Kumar CS, Haritha H, Anand Kumar A, Gopinath S (2018) Feature normalization for improving the performance of sleep apnea detection system. In: IEEE India Council international conference (INDICON), India, Dec 2018

Comparative Analysis of LSTM, One-Class SVM, and PCA to Monitor Real-Time Malware Threats Using System Call Sequences and Virtual Machine Introspection

Jayesh Soni, Suresh K. Peddoju, Nagarajan Prabakar, and Himanshu Upadhyay

Abstract System call analysis is based on a behavior-oriented anomaly detection technique, which is well accepted due to its consistent performance. This study compares two popular algorithms long short-term memory (LSTM) sequence to sequence (Seq-Seq), and one-class support vector machines (OCSVM) for anomalous system call sequences detection. The proposed framework monitors running processes to recognize compromised virtual machines in hypervisor-based systems. The evaluated results show the comparative analysis and effectiveness of feature extraction strategies and anomaly detection algorithms based on their high detection accuracy and with a low loss. This study demonstrates a comparative analysis of detecting anomalous behavior in any process using OCSVM and LSTM Seq-Seq algorithms. A bag-of-2-g with PCA feature extraction strategy and LSTM Seq-Seq with a sequence length of five provides higher detection accuracy of 97.2%.

Keywords Virtual machine (VM) · Long short-term memory (LSTM) · Sequence to sequence (seq-seq) · One-class support vector machines (OCSVM) · Virtual machine introspection (VMI)

1 Introduction

Nowadays, virtualization environments are playing a vital role in different application domains. The major advantage of virtualization is the most utilization of resources. Due to virtualization, there is vast resource sharing, and the scope of protecting resources has improved. Advancements in virtualization methods evolved in extracting virtual machine (VM) [1, 1] information and detected numerous malicious activities are a challenging task. At the present-day market, more sophisticated

---

J. Soni (✉) · N. Prabakar
School of Computing and Information Sciences, Florida International University, Miami, USA
e-mail: jsoni@fiu.edu

S. K. Peddoju · H. Upadhyay
Applied Research Center, Florida International University, Miami, USA

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_9

and vigorous malware is spreading where no detection methods are not in a position to identify malicious activities happening on VM. Thus, most of the businesses are facing cyber threats along with data breaching issue. To overcome these problems, hypervisor-enabled technologies are getting popular in detecting malicious activities and the performances of these techniques are far better than guest-based systems. Virtual machine introspection (VMI) is a technique to extract memory insights into VM to analyze cyber threats and detect malicious activities in it [3–5]. VMI is a mechanism to manage run-time activities of VM at the hypervisor level and memory forensics is a suitable method to study VM activities at the memory level.

The main goal of intruders is to run malicious code on the VM to modify the system files and footprints of resources. There are many more security techniques are available in the market to avoid attacks, such as anti-viruses and security patches. Despite these methods, there are various unknown attacks on the VM by anti-virus software and updating security patches. Prospective malware assaults are identified with the study of characteristics of executing software, which is also known as behavioral analysis [6, 7]. Analysis of system call [8, 9] is coming under the same category with which we can identify the malicious behavior of the software.

The anomaly detection methods and feature extraction techniques are compared and implemented using system call sequences. This study examines and compares OCSVM and LSTM Seq-Seq algorithms in detecting anomalous behavior in any process. The detection system generalizes and performs well in detecting previously unseen anomalous data under different machine configurations and typical load balancing of work. Further, feature engineering is done by studying various methods, and appropriate system call patterns were identified.

The rest of the paper consists of eight sections. Section 2 discusses latest research occurring on the current approach. The procedure for feature extraction and feature engineering for system calls traces using VMI application is described in Sect. 3. In Sect. 4, an overview of detection algorithms is explained. Section 5 describes the experimental setup. Experimental results and discussions are presented in Sect. 6. Section 7 concludes the proposed work, and finally, Sect. 8 proposes the future scope of the research.

2 Related Work

Host-based malware detection frameworks are devised to detect malware on a host system by capturing and studying system-related access. These host-based frameworks utilize signatures of existing malware and try to detect with detection algorithms. The limitation of these frameworks is to detect unknown malware footprints. Apart from this, there are some techniques that are based on the behavior of executing the program in its regular mode. There are different data processing techniques [10, 10] that are evolved in recent times to address these issues and automatically learn and train models based on its characteristics such as “good behavior” and “bad behavior” with benign systems [12, 12]. Behavioral-based malware detection methods are used to address the limitations of static analysis based on traditional detectors. Static analysis techniques are outperforming in their detection based on static features, which can be quickly puzzled [14]. Authors in papers [15, 16] discuss different behavior-related analysis methods. Anomaly detection is placing a significant role in analyzing system call data with the help of statistical and signature detectors. They have also considered several ways of extraction features from the document classification mentioned [17, 18].

Neural networks in combination with anomaly detection are a popular and good idea to detect malicious activities [19–24]. A recent study proves that deep learning methods are extensively used in detecting anomalies, and especially LSTM plays a vital role [25–28]. However, their work was a feature-based supervised classifier, which requires much work to create labels. As a result, there is a chance of labels attack, which inherently creates limitations to fail to detect any unknown attacks [29]. Furthermore, this method requires careful and specific feature engineering techniques which will provide meaningful representations in terms of features for the supervised classification problems [30, 31]. To train their model, only one binary label per sequence was provided, unlike our proposed method, which is trained to predict the next few system calls and that effectively captures contextual information. Few other papers [32–34] discussed the different machine learning algorithms to train models and detect anomalies in different fields.

In this paper, a comparative study is made with different anomaly detection algorithms to study and identify anomalies in extracted data. The normal behavior of the system is trained, and test data is compared with it to see unusual and abnormalities within it. This training and test data comparison is made with a well-defined model with the help of deep learning algorithms. This procedure identifies unknown attacks happening on VM by extracting data with the VMI technique.

3 System Call Traces with VMI

The proposed approach analyzes system calls extracted by using VMI-based techniques [35, 35] to detect malicious activities happening on VM. The main objective of this technique is to extract system call data of active program under execution and inspects the abnormal behavior. This approach alternates existing defenses such as firewalls, antivirus, and many more in identifying compromised VMs.

The framework shown in Fig. 1 extracts system call traces with the VMI method. Guest VM is monitored by the host (monitoring) system with the help of VMI application programming interfaces such as introspector and security agent modules. The introspector module will collect system call information via the security agent. Then system call traces are analyzed with appropriate detection algorithms to identify and detect malicious behavior of processes. Custom test vectors are created for performance evaluation. System call trace sequences are obtained from Guest VM running Windows 10 operating system.

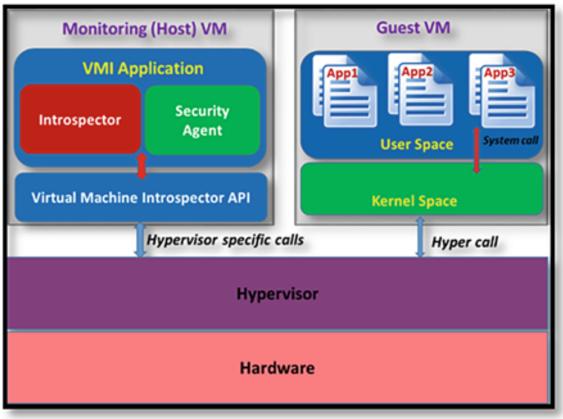


Fig. 1 Virtual machine introspection

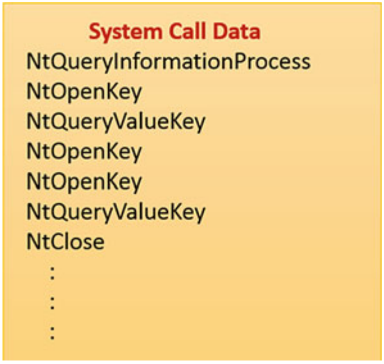

System Call Data
NtQueryInformationProcess
NtOpenKey
NtQueryValueKey
NtOpenKey
NtOpenKey
NtQueryValueKey
NtClose
:
:
:

Fig. 2 System call trace example

A hypervisor is used to extract the system call sequences using a custom VMI application, which includes introspector and agent. This is a service application that receives and logs system call sequences at the process level in recent versions of Microsoft Windows. Figure 2 shows a representation of system call traces, depicting the system calls that occurred during a process. For this study, the VMI application has recorded up to 1,000,000 calls made by each process.

3.1 Feature Engineering for System Call Sequences

In this study, feature engineering is done on system call traces data by using a bag-of-words model. In natural language processing, a document is featured using bag-of-words representation by creating a vector of frequencies of each word. Similarly, the same approach is used by featuring system call sequences as a vector of unique n-gram system call frequencies. An n-gram is a vector of length n of system call sequence

Table 1 Bag-of-2-g representation of system call trace

2-g	Frequency
NtQueryInformationProcess, NtOpenKey	1
NtOpenKey, NtQueryValueKey	2
NtQueryValueKey, NtOpenKey	1
NtOpenKey, NtOpenKey	1
NtQueryValueKey, NtClose	1

occurring continuously in a process under execution. Bag-of-2-g highlighting of the system call trace of Fig. 2 and also is depicted in Table 1. This VMI approach will monitor 450 distinct system calls of windows application.

3.2 Feature Selection

Since features generated through 2-g are sparser, the PCA is utilized to select features with high variance. Orthogonal transformation is used in this method to extract sets of highly correlated variables. Principal components (PCs) are ordered so that the first PC has the most substantial highest variance. Based on PCA, only the primary few components are selected. As a result, the dimension of the dataset is reduced. It is a statistical approach that transforms a set of inter-correlated variables into low dimensions, and such transformed dimensions have a more significant amount of the variability.

4 Detection Algorithms

The anomaly detection learning algorithms are studied and analyzed in the proposed work that operates well for large-scale data. The algorithms exhibit well with less run-time computational complexity during detection. It is an unsupervised learning context, where no label is required in training data for model creation. This context gives an overview of LSTM Seq-Seq and OCSVM.

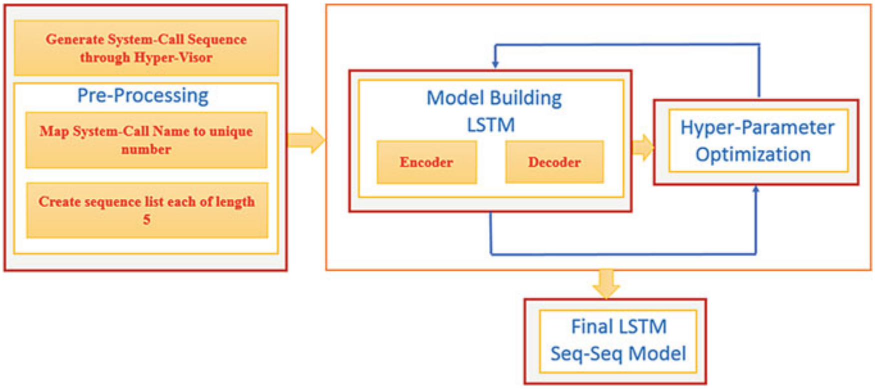

Generate System-Call Sequence through Hyper-Visor
Pre-Processing
Map System-Call Name to unique number
Create sequence list each of length 5
Model Building LSTM
Encoder
Decoder
Hyper-Parameter Optimization
Final LSTM Seq-Seq Model

Fig. 3 Preprocessing and LSTM Seq-Seq model building mechanism

4.1 Long Short-Term Memory (LSTM) Sequence to Sequence (Seq-Seq)

4.1.1 Overview of Our Approach

The LSTM Seq-Seq model is a model based on system calls considering its sequence. Figure 3 depicts our proposed method. There are two parts to our approach, the first is the preprocess data, and the second is the model using LSTM by tuning hyperparameter for efficient optimizations.

4.1.2 Prediction Model

Consider the system call sequences generated by hypervisor during program execution as sentences. Natural language processing (NLP)-based sequence-sequence architecture by feeding the first few system call sequences as the input, which in turn generates the next sequence of system call as output. The sequence to sequence architecture is an encoder–decoder framework based on RNN [37], and it can be viewed as a sentence-to-sentence mapping. The idea of encoder and decoder works in the way that humans think. The same level of processing can be used in system call sequences, consideration is based on the first part as a source sequence in a Q&A system, and the second part is based on the answer.

To transform each document into its numerical vector, this creates its vocabulary. Similarly, a unique lexicon for each system call for a windows machine is employed. Let us assume that there are m individual system calls. A set of all system calls are defined by T, and these are extracted by hypervisor while executing a program. Convert each system call to a numeric vector. The system call sequence s can be represented as $s = (s_1, s_2, s_3, \ldots, s_n)$, and the target system call sequence as $t = (t_1, t_2, t_3, \ldots, t_m)$, where s_i, $t_i \in T$.

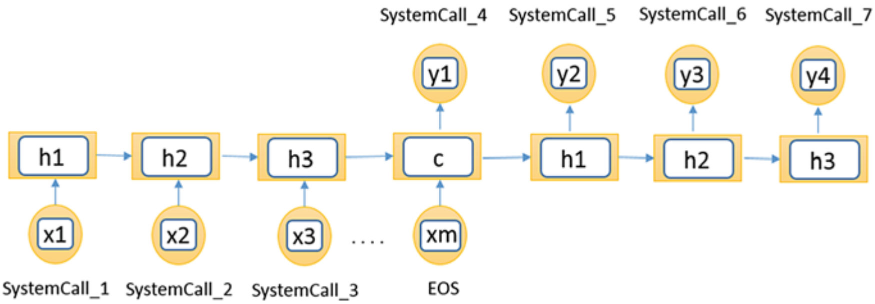


Fig. 4 LSTM encoder–decoder architecture

Given a system call sequence, the encoder produces two states (hidden and output) by completing a forward propagation operation. The formula is as follows:

$$ht = (\mathrm{c}^{HX} x_t + \mathrm{c}^{HH} h_{t-1}) \tag{1}$$

$$y_t = \mathrm{c}^{YH} h_t \tag{2}$$

h_t is encoded information vector c as shown in Fig. 4.

The weight update equations for an LSTM cell are as follows:

$$y_t = \sigma(T_{xi}x_t + T_{hi}k_{t-1} + b_i) \tag{3}$$

$$f_t = \sigma\left(T_{xf}x_t + T_{hf}k_{t-1} + b_f\right) \tag{4}$$

$$o_t = \sigma(T_{xo}x_t + T_{ho}k_{t-1} + b_o) \tag{5}$$

$$g_t = \tanh(T_{xc}x_t + T_{hc}k_{t-1} + b_c) \tag{6}$$

$$c_t = f_t c_{t-1} + i_t g_t \tag{7}$$

$$k_t = o_t \tanh(c_t) \tag{8}$$

In above equations, **tan h** is the hyperbolic tangent function, $\boldsymbol{\sigma}$ is a sigmoid function, and $\mathbf{i_t}$, $\mathbf{f_t}$, $\mathbf{o_t}$, and $\mathbf{c_t}$ are the input gate, forget gate, output gate, and memory cell activation vectors, respectively.

Using context vector c and initial hidden state, the conditional probability of the decoder is shown in Eq. 9.

$$p(y_1, \ldots, y_{T'} | x_1, \ldots, x_1) = \prod_{t=1}^{T'} p(y_t | c, y_1, \ldots, y_{t-1}) \tag{9}$$

Furthermore, perform the attention mechanism by which the conditional probability changes to the below equation.

$$p(y_i | y_1, \ldots, y_{i-1}, X) = g\left(y_{i-1}, s_{i,} c_i\right) \tag{10}$$

where c_i is the context vector calculated during training:

$$c_i = \sum_{j=1}^{T_x} a_{ij} h_j \tag{11}$$

where a_{ij} coefficient of hidden state h_j at time step j. a_{ij} is calculated as follows:

$$a_{ij} = \frac{\exp\left(e_{ij}\right)}{\sum_{k=1}^{T_x} \exp(e_{ik})} \tag{12}$$

where e_{ij} is the relationship learned during the training phase between input and output.

The system call sequence analysis is performed by varying sequence lengths. Finally, 1,000,000 training sequences are collected.

4.2 One-Class Support Vector Machine (OCSVM)

Notational, let us say dataset $\{x_1, x_2, \ldots, x_\mathrm{i}, \ldots, x_\mathrm{N}\}$, with each $X_i \in R^D$ is the one class. The motive of the OCSVM [38] is to detect a hyperplane that separates data from the origin with a high margin. The hyperplane is in high-dimensional space for nonlinear problems with nonlinear transformation $\Phi(.)$. The following quadratic function is used:

$$f(x) = \mathrm{sgn}((w.\emptyset(x)) - \rho) \tag{13}$$

where w is a perpendicular vector to the maximizing hyperplane, and ρ is the distance. A set of variables $\xi i \geq 0$ is introduced to detect outliers. Further, the following decision functions are used.

The OCSVM is trained with and without principal component analysis using system call sequences.

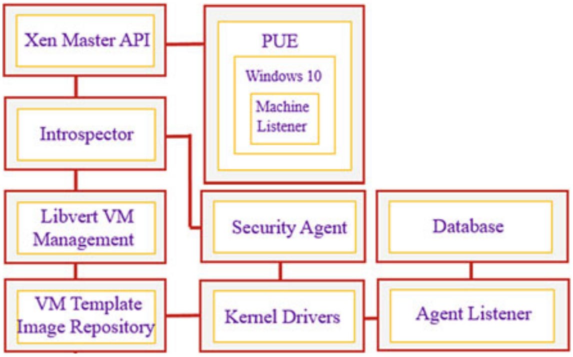


Fig. 5 Framework of proposed evaluation method

5 Experimental Setup

System call data is extracted through a virtual machine, with the help of the Xen hypervisor and libvirt library. Virtual memory introspection is done through DRAKVUF. There are two divisions named introspector and security agent developed using GO language, for data extraction. The virtual machine under inspection is called the system under test (SUT). The recall profile of Google is used for extracting the data. It is a JSON file which contains the memory mappings-related resources. LibVMI library services introspection requests. Various operations of VMs are handled using the libvirt library. An application is designed for performing virtual machine operations and further extracts data. The application is written in the Microsoft Visual Studio.NET framework comprised of user-defined API calls for introspector communication and other related function calls. With the help of an agent, the extracted data is stored into the database. Furthermore, extensive analysis using deep learning algorithms is performed on this data to gain relevant insights. Figure 5 shows the experimental setup.

6 Experimental Results and Discussion

In this subsection, the results of detection algorithms were discussed.

6.1 LSTM Seq-Seq Model Results

In the training of our proposed LSTM Seq-Seq model, the tuning of hyperparameters is critical. The hyperparameters such as sequence length, number of nodes, dropout rate, number of epochs, and batch size are incorporated.

Accuracy as a metric for comparing both the models is defined as follows:

$$\text{Accuracy} = \frac{\text{\# correctly predicted sequences}}{\text{Total sequences of system calls}} \quad (14)$$

The LSTM Seq-Seq neural network in Python using Keras is implemented as the framework. The model with different sequence length values is tuned properly. The dropout rate is 0.2, batch size as 128, and trained with 100 epochs. Different sequence length applied is {3, 5, 10}. The LSTM model is trained with 1,000,000 records and optimized using Adam optimizer. Figures 6 and 7 depict that the sequence length is five and also the low loss and high accuracy, respectively.

Optimizing the number of nodes is critical since it directly affects training time. Utilization of the #Nodes: {16, 32, 64, 128, and 256} is for keeping the other parameters constant.

Figures 8 and 9 depict that 64 nodes perform well in comparison with the varying number of nodes and further increasing the nodes results in model overfits. This model is able to detect the next sequence of system call with high accuracy of 97.2%. On checking against malicious sequences, the model is correctly able to detect malicious behavior.

Fig. 6 Loss versus epoch

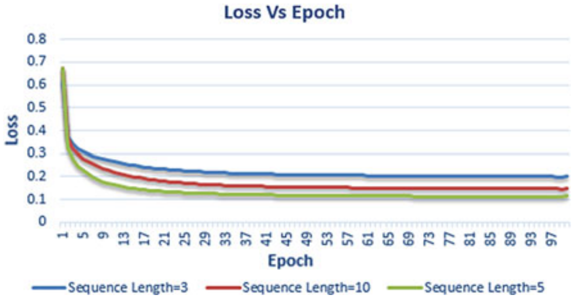


Fig. 7 Epoch versus accuracy

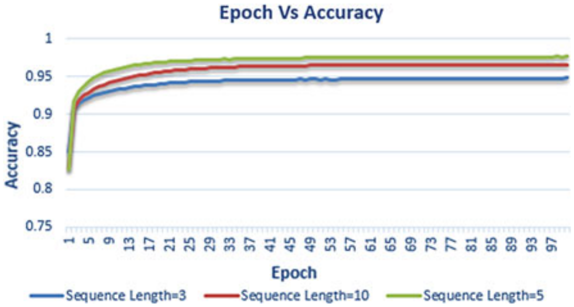


Fig. 8 Training versus testing accuracy

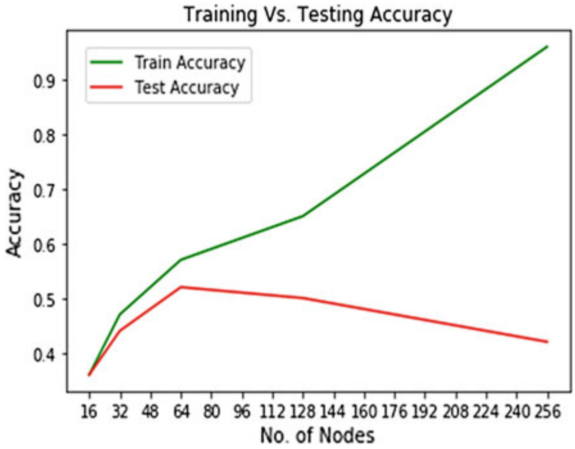
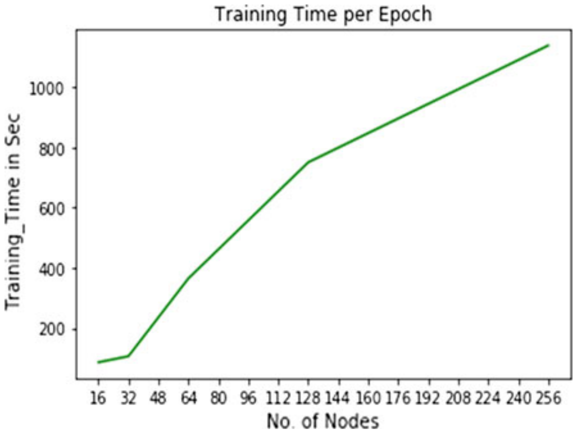


Fig. 9 Training time per epoch w.r.t number of nodes



6.2 *OCSVM Model Results*

For evaluating the OCSVM technique, the similarity score is a metric. The similarity score is a measure which checks how much the training data matches with the testing data. A high similarity score with benign sequences and low similarity scores with malicious sequences is a good sign for an accurate model.

Figures 10 and 11 depict the results of OCSVM trained with PCA on benign sequences. The similarity score with another benign sequence is 90.7% as seen in Fig. 10, whereas the similarity score with malicious sequences is 2.56% as found in Fig. 11.

Furthermore, the OCSVM is trained with raw features without PCA components. The similarity score with another benign sequence is 87.2% as seen in Fig. 12, whereas the similarity score with malicious sequences is 85.56% as found in Fig. 13.

Fig. 10 OCSVM with PCA on benign sequences

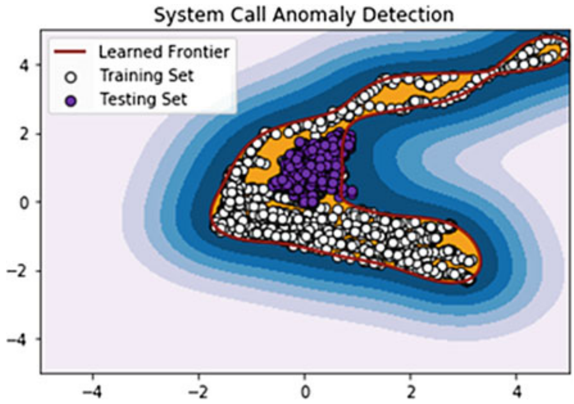


Fig. 11 OCSVM with PCA on malicious sequences

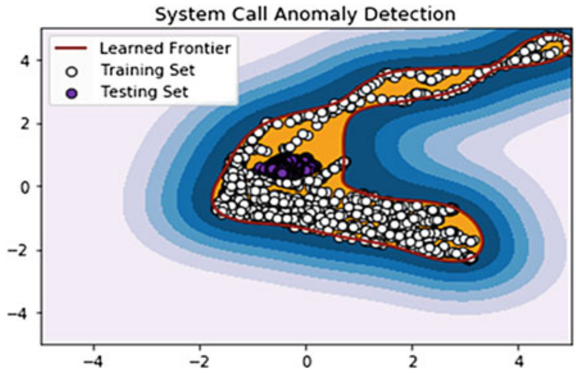


Fig. 12 OCSVM without PCA on benign sequences

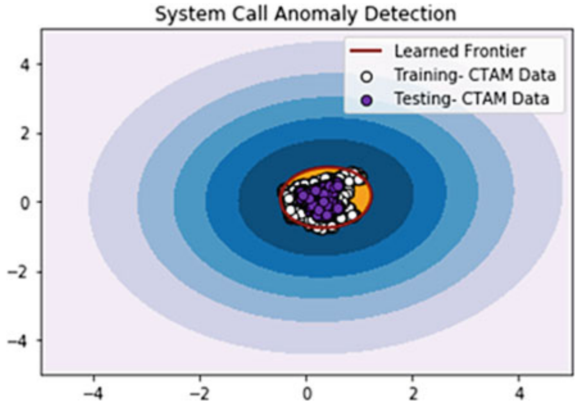


Fig. 13 OCSVM without PCA on benign sequences

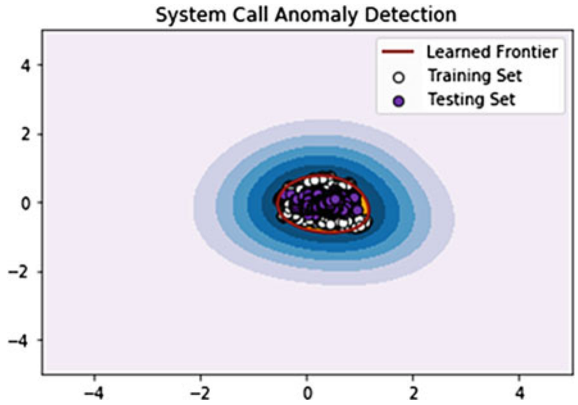


Table 2 Comparative results of the detection algorithms

Model	Accuracy (%)
LSTM Seq-Seq	97.2
OCSVM with PCA	91.6
OCSVM without PCA	87.4

Table 2 lists the testing results of LSTM Seq-Seq, OCSVM with PCA, and OCSVM without PCA on the test set and observed that the LSTM Seq-Seq has the highest accuracy. Moreover, OCSVM with PCA improves the accuracy of a small amount compared with OCSVM without PCA, which indicates that the training on PCA components features is better than the raw features on anomaly detection.

7 Conclusion

The comparative analyses of anomaly detection algorithms, namely LSTM Seq-Seq and one-class SVM with and without PCA using system call sequences data structure are discussed in the proposed work. The introspection technique is used to extract relevant data structures. Subsequently, the filtering and ordering techniques are applied to discard the redundant system calls and obtain valid sequences. To experiment, 1,000,000 samples as datasets are collected. Through the experiment, it is achieved that the LSTM Seq-Seq has the best performance for anomaly detection. It performed with an accuracy of 97.2% as compared to OCSVM. Furthermore, OCSVM with PCA gives better results than OCSVM without PCA.

8 Future Scope

The LSTM Seq-Seq and one-class SVM are applied with and without PCA for anomaly detection using system call sequences. Further, extend this work by applying new natural language processing-based algorithms like transformers, BERT, and variants of LSTM such as bi-directional LSTM and GRU. Furthermore for preprocessing, apart from PCA apply the LDA and autoencoder for feature extraction.

Acknowledgements This work is supported by TRMC, USA.

References

1. Peddoju SK, Upadhyay H, Lagos L (2020) File integrity monitoring tools: issues, challenges, and solutions. Concurr Comput Pract Exper e5825. https://doi.org/https://doi.org/10.1002/cpe.5825
2. Suresh Kumar P, Ramachandram S (2019) Fuzzy-based integration of security and trust in distributed computing. In: Soft computing for problem solving. Advances in intelligent systems and computing, vol 816. Springer, Singapore
3. Ligh MH, Case A, Levy J, Walters A (2014) The art of memory forensics
4. Xen Project (2013) available at https://www.xenproject.org/
5. Hizver J, Chiueh T (2014) Real-time deep virtual machine introspection and its applications.ACM SIGPLAN Notices 49(7)
6. Egele M, Scholte T, Kirda E, Kruegel C (2008) A survey on automated dynamic malware-analysis techniques and tools. ACM Comput Surv
7. Suresh Kumar P, Ramachandram S (2017)Fuzzy based integration of security and trust in distributed computing. In: Proceedings of Springer 7th international conference soft computing for problem solving (SocProS'2017). Indian Institute of Technology, Bhubaneswar, December 2017
8. Forrest S, Hofmeyr S, Somayaji A, Longstaff T (1996) A sense of self for UNIX processes. In: IEEE Security and Privacy
9. Joshi N, Choksi DB (2014) Implementation of process forensic for system calls. Int J Adv Res Eng Technol (IJARET) 5(6):77–82. ISSN 0976–6480(Print), ISSN 0976–6499
10. Lee W, Stolfo S, Mok K (1999) A data mining framework for building intrusion detection models. In: IEEE symposium on security and privacy, pp 120–132
11. Mahoney M, Chan P (2001) Detecting novel attacks by identifying anomalous network packet headers. Technical Report CS-2001-2
12. Suresh Kumar P, Pranavi S (2017) Performance analysis of machine learning algorithms on diabetes dataset using big data analytics. In: Proceedings of IEEE 2017 international conference on infocom technologies and unmanned systems (ICTUS'2017), Dubai, United Arab Emirates (UAE), December 2017, pp 580–585
13. Rishika Reddy A, Suresh Kumar P (2016) Predictive big data analytics in healthcare. In: Proceedings of IEEE 2016 second international conference on computational intelligence & communication technology (CICT), Ghaziabad, pp 623–626
14. Moser A, Kruegel C, Kirda E (2007) Limits of static analysis for malware detection. In: Annual computer security applications conference, ACSAC
15. Ye N, Li X, Chen Q, Emran SM, Xu M (2001) Probabilistic techniques for intrusion detection based on computer audit data. In: IEEE transactions on systems, man, and cybernetics
16. Zhang Q, Reeves D (2007) Metaaware: identifying metamorphic malware. In: Annual computer security applications conference.

17. Kang D-K, Fuller D, Honavar V (2005) Learning classifiers for misuse and anomaly detection using a bag of system calls representation. In: Annual information assurance workshop
18. Xiao H, Stibor T (2011) A supervised topic transition model for detecting malicious system call sequences. In: The workshop on knowledge discovery, modeling, and simulation
19. Debar H, Becker M, Siboni D (1992) A neural network component for an intrusion detection system. In: Proceedings of 1992 IEEE computer society symposium on research in security and privacy, pp 240–250. IEEE
20. Ryan J, Lin M-J, Miikkulainen R (1998) Intrusion detection with neural networks. Adv Neural Inform Proces Syst 943–949
21. Mukkamala S, Janoski G, Sung A (2002) Intrusion detection using neural networks and support vector machines. In: Proceedings of the 2002 international joint conference on neural networks, 2002. IJCNN'02, vol 2, pp 1702–1707. IEEE
22. Wang G, Hao J, Ma J, Huang L (2010) A new approach to intrusion detection using artificial neural networks and fuzzy clustering. Expert Syst Appl 37(9):6225–6232
23. Creech G, Jiankun Hu (2014) A semantic approach to host-based intrusion detection systems using contiguous and discontiguous system call patterns. IEEE Trans Comput 63(4):807–819
24. Suresh Kumar P, Upadhyay H, Bansali S (2019) Health monitoring with low power IoT devices using anomaly detection algorithm. In: IEEE conference SLICE-2019, Rome, Italy
25. Staudemeyer RC (2015) Applying long short-term memory recurrent neural networks to intrusion detection. South African Comput J 56(1):136–154
26. Staudemeyer RC, Omlin CW (2013) Evaluating performance of long short-term memory recurrent neural networks on intrusion detection data. In: Proceedings of the South African institute for computer scientists and information technologists conference, pp 218–224. ACM
27. Soni J, Prabakar N, Upadhyay H (2019) Deep learning approach to detect malicious attacks at system level. In: WiSec'19: Proceedings of 12th ACM conference on security & privacy in wireless and mobile networks, May 15–17, 2019, Miami, FL, USA, 2p
28. Soni J, Prabakar N (2018) Effective machine learning approach to detect groups of fake reviewers. In: Proceedings of the 14th international conference on data science (ICDATA'18), Las Vegas, NV
29. Soni J, Prabakar N, Upadhyay H (2019) Feature extraction through deepwalk on weighted graph. In: Proceedings of the 15th international conference on data science (ICDATA'19), Las Vegas, NV
30. Soni J, Prabakar N, Kim J-H (2017) Prediction of component failures of telepresence robot with temporal data. In: 30th Florida conference on recent advances in robotics
31. Thejas GS, Soni J, Chandna K, Iyengar SS, Sunitha NR, Prabakar N (2019) Learning-based model to fight against fake like clicks on Instagram posts. In: SoutheastCon 2019. Huntsville, Alabama, USA. In Press
32. Tejaswi U, Suresh Kumar P (2016) Diagnosing diabetes using data mining techniques. Int J Sci Res Publ 7(6):705–709
33. Chaudhary A, Peddoju SK, Peddoju SK (2020) Cloud based wireless infrastructure for health monitoring. Virt Mob Healthcare 34–55
34. Peddoju SK, Kavitha K, Sharma SC (2019) Big data analytics for childhood pneumonia monitoring. In: IGI global edited book, 2019, pp 1–17
35. Peddoju SK, Upadhyay H (2020) Evaluation of IoT data visualization tools and techniques. Data visualization. Springer, Berlin
36. Peddoju SK, Upadhyay H, Soni J, Prabakar N (2020) Natural language processing based anomalous system call sequences detection with virtual memory introspection. Int J Adv Comput Sci Appl (IJACSA) 11(5). http://dx.doi.org/https://doi.org/10.14569/IJACSA.2020.0110559
37. Bahdanau D, Cho K, Bengio Y (2014) Neural machine translation by jointly learning to align and translate. Comput Sci
38. Schölkopf B, Platt JC, Shawe-Taylor J, Smola AJ, Williamson RC (2001)Estimating the support of a high-dimensional distribution. Neural Comput 13(7):1443–1471

SBP: Preparation of Schema Elements to Schema Matching

Aola Yousfi, Moulay Hafid El Yazidi, and Ahmed Zellou

Abstract Schema pre-matching is very critical for having schema elements fully ready for schema matching. Ideally, words are first extracted from the schema elements' labels, and then the semantically corresponding elements are generated accordingly. Searching for the sense of words based on their vertical and horizontal contexts, and before performing schema matching is very crucial to obtain high matching accuracy and as a result increase the amount of accurate matches and reduce the number of inaccurate matches and missed matches. Nonetheless, this problem is much more challenging than it seems. This is because complete and precise information about the meaning behind each element is often unavailable. This paper presents SBP, a Sense-Based Pre-matching approach designed for hierarchical data representations. SBP consists of two main components. First, the words sets generator generates, from each element, a set of words that fully describes its meaning. Second, the words qualifier identifies the senses of words based on both their vertical and horizontal contexts. Experimental results on real-world domains show high matching accuracy obtained when using the SBP approach.

Keywords Schema matching · Schema pre-matching · Semantic similarity · Vertical context · Horizontal context · Matching accuracy

---

A. Yousfi (✉) · M. H. El Yazidi · A. Zellou
Software Project Management Research Team, ENSIAS,
Mohammed V University, Rabat, Morocco
e-mail: aola.yousfi@gmail.com

M. H. El Yazidi
e-mail: my-hafid.elyazidi@um5.ac.ma

A. Zellou
e-mail: ahmed.zellou@um5.ac.ma

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_10

1 Introduction

Schema matching is very critical for applications that manipulate data across schemas of distinct data sources, examples of areas where this kind of applications are used include mainly data integration on the World Wide Web, data warehousing, e-commerce, scientific collaboration and bioinformatics. Schema matching not only requires a semantic comparison between elements from different schemas, but also needs an identification of the full meaning of each schema element before proceeding with the matching, which is called schema pre-matching or schema pre-processing. This paper presents two interesting observations for schema pre-processing. First, the labels of the schema elements are ambiguous. They often include acronyms (that correspond to the acronyms and abbreviations database entries), abbreviations (that also correspond to the acronyms & abbreviations database entries) and words (that correspond to the lexical dictionary entries) separated by underscores (e.g. *academic_conf_name*) or juxtaposed against each other (e.g. *academicConfName*). Second, the meanings, also known as senses, of words often change in different contexts.

Let S_1 and S_2 bet two schemas, and let e_1 and e_2 be two semantically corresponding elements (also called matches according to [1]) from S_1 and S_2, respectively. If e_1 and e_2 happens to use the same exact naming convention, then schema matching would be straightforward and pre-matching would not be a topic of discussion. Nevertheless, since there is no universal naming standard, schema pre-matching is very critical when matching new schemas. Therefore, plenty of schema matching systems have been introduced throughout the years (see [2–5] for recent surveys) to search for the matches in different schemas with the idea that semantically corresponding elements may likely be spelled differently. Although the state of the art schema matching systems obtain accurate results, they also obtain inaccurate matches and miss some accurate matches. As a result, these systems will remain completely dependent on human assistance in order to correct the output of the matching systems.

When matching a huge number of schemas, it is way much better to try to get as many accurate matches as possible right from the beginning. Hence, capturing the correct and complete meaning of schema elements prior to generating the semantically corresponding elements is very important to increase the total number of accurate matches and decrease the total number of inaccurate matches and missed matches. The problem of searching for the meanings of schema elements is not easy though.

Next, this paper presents an example that shows the importance of identifying words' senses prior to performing any schema matching operation.

Example 1.1 Let S_1 (see Listing 10.1) and S_2 (see Listing 10.2) be two snippets of two XML schemas describing the domain of organising conferences.

If the matching approach ignores schema elements contexts, then it will end up matching S_1.conference.rented_products.chairs to S_2.conference.organizing_committee.chair (which are not semantically similar). Nonetheless, if the matching approach takes into consideration schema elements contexts {conference, rented_products} for S_1.chairs and {conference, organizing_committee} for S_2.chair, then it will end up not matching S_1.conference.rented_products.chairs to S_2.conference.organizing_committee.chair (simply because rented_products and organizing_committee refer to two completely different real-world entities, which implies that S_1.chair and S_2.chair are in fact homonyms).

Listing 10.1 S_1

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
<xs:element name="conference">
 <xs:complexType>
  <xs:element name="title" type="xs:string"/>
  <xs:element name="date" type="xs:date"/>
  <xs:element name="address" type="xs:string"/>
  <xs:element name="rented_products">
   <xs:complexType>
    <xs:element name="chairs" type="xs:integer"/>
    <xs:element name="tables" type="xs:integer"/>
   </xs:complexType>
   </xs:element>
 </xs:complexType>
</xs:element>
</xs:schema>
```

Listing 10.2 S_2

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
<xs:element name="conference">
 <xs:complexType>
  <xs:element name="name" type="xs:string"/>
  <xs:element name="location" type="xs:string"/>
  <xs:element name="date" type="xs:date"/>
  <xs:element name="organizing_committee">
   <xs:complexType>
    <xs:element name="chair" type="xs:string"/>
    <xs:element name="program_committee" type="xs:string"/>
    <xs:element name="steering_committee" type="xs:string"/>
    <xs:element name="publicity_committee" type="xs:string"/>
   </xs:complexType>
  </xs:element>
 </xs:complexType>
</xs:element>
</xs:schema>
```

This paper introduces SBP, a Sense-Based Pre-matching approach designed for hierarchical data representations. SBP main idea is to prepare schema elements for the schema matching step. It suggests to generate, from each schema element, a words set that fully describes its sense. SBP deals with the problem of imprecise and sometimes unavailable information about the meanings of words as follows: it exploits schema elements' labels, and uses the horizontal and vertical contexts (see Definitions 3.3 and 3.4) of schema elements.

In summary, this paper makes the following concrete contributions:

- It introduces a new approach to capture the complete and correct meaning of schema elements.
- It proposes a novel algorithm that generates, from each schema element, a set of words that fully describes its meaning.

- It proposes a novel equation that determines the accurate sense of a word according to its horizontal and vertical contexts.
- It evaluates SBP on eight real-world domains and show that it significantly improves the matching accuracy: increases the total number of accurate matches, and decreases the total number of missed and inaccurate matches.

The remaining of this paper is structured as follows. Section 2 discusses related work. Section 3 defines the problem of schema pre-matching. Section 4 describes the architecture of SBP. Section 5 presents the experimental results. Section 6 concludes this paper and discusses future research directions.

2 Related Work

To the best of our knowledge, schema pre-matching has received very little attention from the research community compared to schema matching (see [2, 6–9] for surveys on schema matching and [10–17] for current matching tools). Moreover, current schema matching tools proceed directly with the matching as they start searching for semantically corresponding elements between schemas right from the beginning, and do not propose a solution to the pre-matching problem. In what follows, the paper will describe such state of the art matching systems.

PORSCHE [10] (Performance Oriented SCHEma mediation) is an automatic schema matching tool. It uses external thesauri and tree mining techniques, along with string-based information and language-based information. PORSCHE proceeds in three main steps. First, it transforms schemas into rooted ordered labeled trees. Second, it combines all schema trees into one single integrated schema. Third, PORSCHE captures the semantically corresponding elements between the integrated schema and the schema trees.

AgreementMakerLight (AML) [12] derives from AgreementMaker [18]. AML consists of two main modules: ontology loading module and ontology matching module. The ontology loading module loads the ontology files and then uses dictionaries to generate ontology objects. Note that the ontology loading module allows the virtual integration of new matching algorithms. The ontology matching module then aligns the ontology objects generated by ontology loading module.

COMA++ (COmbining Match Algorithms) [19, 20] is an updated version of COMA [19]. It transforms schemas into rooted directed acyclic graphs by means of external dictionaries as well as structure-based information, language-based information and instance-based information.

This section showed that none of the state of the art schema matching tools that are most relevant to our work takes into consideration the different senses a word might have in different contexts, and proposes a pre-processing strategy accordingly. Also, none of the state of the art schema matching tools assigns words sets to schema elements, and identifies the sense of words according to both their vertical and horizontal contexts.

3 Problem Statement

This section presents the definitions used throughout this paper, and states clearly the problem statement.

Definition 3.1 (*Schema element*) Let S be a schema. A schema element e refers to an element from S. Note that a schema element refers to both simple type elements and complex type elements. For example, in S_1 (see Listing 10.1), conference is a schema element and title is also a schema element.

Remark 1 In hierarchical data structures, this paper refers to the leaves by simple type elements, and the inner nodes as well as the root by complex type elements.

Definition 3.2 (*Words Set*) Let S be a schema and e an element from S. A words set θ generated from e is defined as follows: $\theta = \{w_1, w_2, \ldots, w_{|\theta|}\}$, where $\forall i \in [\![1, |\theta|]\!]$, w_i is a word from the lexical dictionary or the acronyms & abbreviations database. For example, the words set generated from chair in S_2 (see Listing 10.2) is defined as follows: $\theta_{\text{chair}} = \{conference, organizing, committee, chair\}$ (see Sect. 4 for more details on the words sets generation process).

Definition 3.3 (*Vertical context*) Let S be a schema and e be an element from S. The vertical context of e refers to all the complex type elements e is contained in. For example, in S_1 (see Listing 10.1), the vertical context of *chairs* is defined as follows: $\theta_{\text{VC}} = \{conference, rented, products\}$.

Definition 3.4 (*Horizontal context*) Let S be a schema and e be an element from S. The horizontal context of e refers to the elements surrounding e in S. For example, in S_1 (see Listing 10.1), the horizontal context of *chairs* is defined as follows: $\theta_{\text{HC}} = \{title, date, address, tables\}$.

Definition 3.5 (*Problem Statement*) Given a schema S, let ce be a complex type element from S and se be a simple type element contained in ce. Our main objective is to exploit the relationship between se and ce, apply some extracting techniques, and use external resources (a hierarchical lexical dictionary along with an acronyms and abbreviations database) to find out a words set that fully describes the meaning of se.

Table 1 lists the notations used throughout this paper.

The next section describes SBP, the solution to the problem presented in Definition 3.5.

Table 1 Summary of symbol notations

Notation	Description
$\mathbb{S}$, S, e	All input schemas, a schema from $\mathbb{S}$, an element from S
θ', $\lvert\theta'\rvert$, Θ', $\mathbb{T}'$	Words set generated from e, cardinality of θ', the sets of words generated from S, the sets of words generated from $\mathbb{S}$
θ, $\lvert\theta\rvert$, Θ, $\mathbb{T}$	Words set generated from e (with the senses of words identified), cardinality of θ, the sets of words generated from S (with the senses of words identified), the sets of words generated from $\mathbb{S}$ (with the senses of words identified)
w	Word
$\theta_{\mathbb{LD}}$	Set of words that corresponds to the lexical dictionary entries
θ_{acr}, θ_{abbr}	Set of words whose acronyms correspond to the acronyms & abbreviations database entries, set of words whose abbreviations correspond to the acronyms and abbreviations database entries
θ_{VC}, θ_{HC}	Vertical context, horizontal context
$\mathbb{LD}$, $DB_{\text{acr\&abbr}}$	Lexical dictionary, acronyms and abbreviations database

4 SBP: The Sense-Based Pre-matching

This section describes the solution to the schema pre-matching problem. The vast majority of current schema matching systems consider mainly hierarchical data structures, such as XML schemas. Hence, this paper focuses on the problem of preparing such schemas for the actual matching, leaving other data representations as future work.

The architecture of the Sense-Based Pre-matching approach (SBP) (see Fig. 1) includes two key components: words sets generator and words qualifier. Let $\mathbb{S}$ be a set of input schemas to match, and let $S \in \mathbb{S}$ be a schema, the *words sets generator* first applies fuzzy string matching techniques, and exploits a hierarchical lexical dictionary $\mathbb{LD}$ along with an acronyms & abbreviations database $DB_{\text{acr\&abbr}}$ so as to generate, from every schema element $e \in S$, a words set $\theta' \in \Theta' \in \mathbb{T}'$ that describes its meaning (Θ' refers to all the words sets generated from S, and $\mathbb{T}'$ refers to all the sets of words generated from $\mathbb{S}$). Then, for every θ', the *words qualifier* employs the horizontal context of e to determine the sense of words, therefore updating the words set from θ' to $\theta \in \Theta \in \mathbb{T}$ (Θ refers to all the words sets generated from S, which in addition to the information already given by Θ', Θ states the senses of words in terms of their context; and $\mathbb{T}$ refers to all the words sets generated from $\mathbb{S}$ which specifies the senses of words).

The rest of this section describes thoroughly the words sets generator in Sect. 4.1 and the words qualifier in Sect. 4.2.

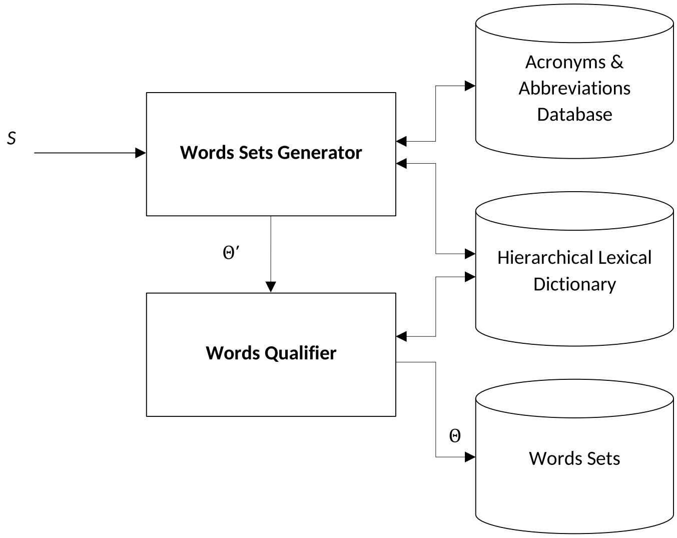

S
Words Sets Generator
Θ′
Words Qualifier
Θ
Acronyms & Abbreviations Database
Hierarchical Lexical Dictionary
Words Sets

Fig. 1 The SBP architecture

4.1 Words Sets Generator

The words sets generator takes as input schemas $S \in \mathbb{S}$ and delivers as output words sets $\Theta' \in \mathbb{T}'$ that do not specify the senses of words just yet. Below, the paper will describe the three steps (summarized in Algorithm 1) the words sets generator goes through before it generates the sets of words.

Step 1: Extract, from each schema element, plain words that correspond to the lexical dictionary entries.
Simple type elements describe the data stored inside data sources. Nonetheless, the simple type elements' labels on their own do not fully describe the definition of the data. Hence, it was decided to use the simple type elements along with the complex type elements they are included in. But, here again a new challenge was faced: the schema elements labels are neither explicit nor complete. Therefore, given a schema element e (e can be a simple type element or a complex type element), the words sets generator uses fuzzy string matching techniques to extract from e words $\theta'_{\mathbb{LD}}$ that correspond to the hierarchical lexical dictionary entries [see formula (1)].

$$e \xrightarrow{\text{convert into}} \theta'_{\mathbb{LD}} \tag{1}$$

Then, the words sets generator assigns $\theta'_{\mathbb{LD}}$ to θ' as presented in formula (2).

$$\theta' \leftarrow \theta' \cup \theta'_{\mathbb{LD}} \tag{2}$$

Step 2: Substitute acronyms and abbreviations with their corresponding full forms.
The words sets generator exploits $DB_{\mathrm{acr\&abbr}}$ to replace acronyms acr in e, if any, with their full forms θ'_{acr} (see formula (3)).

$$\mathrm{acr} \xrightarrow{\text{convert into}} \theta'_{\mathrm{acr}} \tag{3}$$

Then, it assigns the full forms' set θ'_{acr} to θ' as displayed in formula (4).

$$\theta' \leftarrow \theta' \cup \theta'_{\mathrm{acr}} \tag{4}$$

Similarly, the words sets generator uses an acronyms & abbreviations database to replace abbreviations abbr in e, if any, with their full forms θ'_{abbr}, as presented in formula (5).

$$\mathrm{abbr} \xrightarrow{\text{convert into}} \theta'_{\mathrm{abbr}} \tag{5}$$

Then, it assigns the full forms' set θ'_{abbr} to θ' as shown in formula (6).

$$\theta' \leftarrow \theta' \cup \theta'_{\mathrm{abbr}} \tag{6}$$

Step 3: Generate words sets.
Finally, the words sets generator enriches the sets assigned to the simple type elements in order to gain new insights [see formula (7)]: let *se* be a simple type element, the words sets generator converts *se* into a union of its words set and the sets of its vertical context θ'_{VC}.

$$\theta'_{se} \leftarrow \theta'_{se} \cup \theta'_{\mathrm{VC}} \tag{7}$$

4.2 *Words Qualifier*

Given a schema element e, let $\theta' = \{w_1, w_2, \ldots, w_{|\theta'|}\} \in \Theta' \in \mathbb{T}'$ be the words set generated from e, $\forall\, i \in [\![1, |\theta'|]\!]$, $\exists j \geq 1$, such that w_i has j different senses. The words qualifier exploits the horizontal context $\theta'_{\mathrm{HC}} = \{w_{\mathrm{HC}_1}, w_{\mathrm{HC}_2}, \ldots, w_{\mathrm{HC}_{|\theta'_{\mathrm{HC}}|}}\}$ of e, as shown in Eq. (8), so as to identify the appropriate sense of w_i in the given context.
$\forall\, i \in [\![1, |\theta'|]\!]$, $\exists j \geq 1$,

Algorithm 1 WordsSetsGenerator($\mathbb{S}$)

```
Input:
  $\mathbb{S}$ : Input schemas
Output:
  $\mathbb{T}'$ : Words sets

for each $S$ in $\mathbb{S}$ do
  for each $se$ in $S$ do
    if $\exists\ w \in se$ and $w \in \mathbb{LD}$ then
      $\theta'_{se} \leftarrow \theta'_{se} \cup w$
    end if
    if $\exists$ abbreviation $abbr \in se$ and $abbr \in DB_{acr\&abbr}$ then
      Replace $abbr$ with its full form
      Add its expanded form to $\theta'_{se}$
    end if
    if $\exists$ acronym $acr \in se$ and $acr \in DB_{acr\&abbr}$ then
      Replace $acr$ with its full form
      Add its expanded form to $\theta'_{se}$
    end if
  end for
  for each $\theta'_{se}$ in $\Theta'$ do
    $\theta'_{se} \leftarrow \theta'_{se} \cup \theta'_{VC}$
  end for
end for
return $\mathbb{T}'$
```

$$
\begin{aligned}
\text{score}_j(w_i)_{1 \le j \le |\text{senses}(w_i)|} = {} & |[\text{Definition}_j(w_i) \cup \text{Synonyms}_j(w_i) \cup \text{Examples}_j(w_i)] \\
& \cap [(\bigcup_{\substack{k=1 \\ \theta' \setminus w_i}}^{|\theta'|} \bigcup_{q=1}^{|\text{senses}(w_k)|} (\text{Definition}_q(w_k) \cup \text{Synonyms}_q(w_k) \\
& \cup \text{Examples}_q(w_k))) \\
& \cup (\bigcup_{o=1}^{|\theta'_{\text{HC}}|} \bigcup_{p=1}^{|\text{senses}(w_{\text{HC}_o})|} (\text{Definition}_p(w_{\text{HC}_o}) \\
& \cup \text{Synonyms}_p(w_{\text{HC}_o}) \cup \text{Examples}_p(w_{\text{HC}_o})))]|
\end{aligned}
\tag{8}
$$

where

- $|\text{senses}(w_i)|$ refers to the total number of senses of w_i in $\mathbb{LD}$.
- $\text{Definition}_j(w_i)$ is the definition of the jth sense of w_i in $\mathbb{LD}$.
- $\text{Synonyms}_j(w_i)$ is the set of synonyms of the jth sense of w_i in $\mathbb{LD}$.
- $\text{Examples}_j(w_i)$ are the examples assigned to the jth sense of w_i in $\mathbb{LD}$.
- w_k denotes a word from $\theta' \setminus w_i$.
- Given a word $w_k \in \theta' \setminus w_i$, $|\text{senses}(w_k)|$ refers to the total number of senses of w_k in $\mathbb{LD}$.
- $\text{Definition}_q(w_k)$ is the definition of the qth sense of w_k in $\mathbb{LD}$.

- $\mathrm{Synonyms}_q(w_k)$ is the set of synonyms of the qth sense of w_k in $\mathbb{LD}$.
- $\mathrm{Examples}_q(w_k)$ are the examples assigned to the qth sense of w_k in $\mathbb{LD}$.
- Given a word $w_{\mathrm{HC}_o} \in \theta'_{\mathrm{HC}}$, $|\mathrm{senses}(w_{\mathrm{HC}_o})|$ refers to the total number of w_{HC_o} senses in $\mathbb{LD}$.
- $\mathrm{Definition}_p(w_{\mathrm{HC}_o})$ is the definition of the pth sense of w_{HC_o} in $\mathbb{LD}$.
- $\mathrm{Synonyms}_p(w_{\mathrm{HC}_o})$ is the set of synonyms of the pth sense of w_{HC_o} in $\mathbb{LD}$.
- $\mathrm{Examples}_p(w_{\mathrm{HC}_o})$ are the examples assigned to the pth sense of w_{HC_o} in $\mathbb{LD}$.

Then, the words qualifier assigns the sense j of w_i with the largest $score_j(w_i)$ to w_i:

$$\begin{aligned}\mathrm{sense}(w_i) \leftarrow \mathrm{sense}_j(w_i), \text{ such that } \mathrm{score}_j(w_i) \\ = \max(\mathrm{score}_j(w_i))_{1\leq j\leq|\mathrm{senses}(w_i)|}\end{aligned} \tag{9}$$

As a consequence, the following results are obtained such that the sense of every single word in θ' is stated clearly.

$$\begin{aligned}\theta' = \{w_1, w_2, \ldots, w_{|\theta'|}\} \xrightarrow{\text{convert into}} \theta \\ = \{w_1\#n\#s, w_2\#n\#s, \ldots, w_{|\theta|}\#n\#s\}\end{aligned} \tag{10}$$

where

- $\#n$ is short for noun.
- $\#s$ is the sth sense of w_i (where $i \in [\![1, |\theta|]\!]$) in $\mathbb{LD}$, such that $s \in \mathbb{N}$. $w_i\#n\#1$ is the first meaning of w_i in $\mathbb{LD}$, $w_i\#n\#2$ is the second meaning of w_i in $\mathbb{LD}$, and so on.

Algorithm 2 summarizes this.

Algorithm 2 WordsQualifier($\mathbb{T}'$)

```
Input:
  $\mathbb{T}'$ : Words sets before sense identification
Output:
  $\mathbb{T}$ : Words sets with the senses identified

for each $\Theta'$ in $\mathbb{T}'$ do
  for each $\theta'$ in $\Theta'$ do
    for each $w$ in $\theta'$ do
      Calculate the score of $w$ according to Eq. (8)
      Keep only the sense with the maximal score
    end for
  end for
end for
return $\mathbb{T}$
```

5 Experimental Results

The paper runs extensive experiments to assess SBP using real implementation. It mainly focuses on evaluating two key issues. First, it examines the accuracy of the generated words sets by comparing them against the reference sets. The reference words sets were found manually by a group of forty-five Ph.D. candidates from our university who also specified the exact sense of words based on the words contexts. Second, it verifies the ability of SBP to determine correct matches by applying it to some current schema matching tools.

5.1 Experimental Settings

Datasets: This paper experimented SBP on eight datasets (see Table 2) from TEL (Travel, Entertainment and Living) which are publically available on the Web.[1] The travel domain groups its dataset into three distinct sub-domains: *Car Rentals*, *Hotels* and *Airfares*. The Entertainment domain groups its dataset into three distinct sub-domains: *Music Records*, *Movies* and *Books*. And, the Living domain groups its datasets into two distinct sub-domains: *Automobiles* and *Jobs*.

Implementation: This paper first implements SBP using WordNet[2] [21] as our hierarchical lexical dictionary, and evaluates the accuracy of the sets of words it generates. Then, it uses SBP with COMA++ [19, 22] (COMA++$_{\text{SBP}}$), PORSCHE [10] (PORSCHE$_{\text{SBP}}$) and AML [9, 12, 18] (AML$_{\text{SBP}}$), and compare the results to those obtained by COMA++, PORSCHE and AML.

Measures: This paper first exploits the metrics [19] defined in (11)–(14) to evaluate the accuracy of the words sets generated by SBP.

$$\text{Precision}_{\text{Sets}} = \frac{\text{Accurate Sets}}{\text{Accurate Sets} + \text{Inaccurate Sets}} \tag{11}$$

(11) identifies the percentage of accurate sets among all sets returned by SBP.

$$\text{Recall}_{\text{Sets}} = \frac{\text{Accurate Sets}}{\text{Missed Sets} + \text{Accurate Sets}} \tag{12}$$

(12) determines the percentage of accurate sets returned by SBP among all reference sets.

$$F\text{-Measure}_{\text{Sets}} = \frac{2 \times \text{Precision}_{\text{Sets}} \times \text{Recall}_{\text{Sets}}}{\text{Precision}_{\text{Sets}} + \text{Recall}_{\text{Sets}}} \tag{13}$$

---

[1] http://metaquerier.cs.uiuc.edu/repository/datasets/tel-8/browsable.html.

[2] http://wordnetweb.princeton.edu/perl/webwn.

Table 2 Evaluation datasets

Domain	Total number of schemas
Car rentals	25
Hotels	39
Airfares	47
Music records	65
Movies	73
Books	65
Automobiles	84
Jobs	49

(13) is the harmonic mean of $\text{Precision}_{\text{Sets}}$ and $\text{Recall}_{\text{Sets}}$.

$$\text{Overall}_{\text{Sets}} = \text{Recall}_{\text{Sets}} \times \left(2 - \frac{1}{\text{Precision}_{\text{Sets}}}\right) \tag{14}$$

(14) determines the manual post-effort required to remove inaccurate sets and add missed sets. Unlike $\text{Precision}_{\text{Sets}}$ and $\text{Recall}_{\text{Sets}}$, $\text{Overall}_{\text{Sets}}$ may have negative values if $\text{Precision}_{\text{Sets}} < 0.5$. Note that if $\text{Overall}_{\text{Sets}} < 0$ then most of the pre-matching work is going to be performed manually.

The ideal case scenario is when $\text{Precision}_{\text{Sets}} = \text{Recall}_{\text{Sets}} = F\text{-Measure}_{\text{Sets}} = \text{Overall}_{\text{Sets}} = 1$.

The paper then employs the metrics [19] defined in (15)–(18) to evaluate the quality of the matches generated by $\text{COMA++}_{\text{SBP}}$, $\text{PORSCHE}_{\text{SBP}}$ and AML_{SBP}.

$$\text{Precision}_{\text{Matches}} = \frac{\text{Accurate Matches}}{\text{Accurate Matches} + \text{Inaccurate Matches}} \tag{15}$$

(15) determines the percentage of accurate matches among all matches generated by the matching tool.

$$\text{Recall}_{\text{Matches}} = \frac{\text{Accurate Matches}}{\text{Missed Matches} + \text{Accurate Matches}} \tag{16}$$

(16) identifies the percentage of accurate matches generated by the matching tool among all reference matches (i.e. matches found manually).

$$F\text{-Measure}_{\text{Matches}} = \frac{2 \times \text{Precision}_{\text{Matches}} \times \text{Recall}_{\text{Matches}}}{\text{Precision}_{\text{Matches}} + \text{Recall}_{\text{Matches}}} \tag{17}$$

(17) is the harmonic mean of $\text{Precision}_{\text{Matches}}$ and $\text{Recall}_{\text{Matches}}$.

$$\text{Overall}_{\text{Matches}} = \text{Recall}_{\text{Matches}} \times \left(2 - \frac{1}{\text{Precision}_{\text{Matches}}}\right) \tag{18}$$

(18) identifies the amount of manual post-effort necessary to remove inaccurate matches and add missed matches. Different from $\text{Precision}_{\text{Matches}}$ and $\text{Recall}_{\text{Matches}}$, $\text{Overall}_{\text{Matches}}$ might have a negative value if $\text{Precision}_{\text{Matches}} < 0.5$. Note that if $\text{Overall}_{\text{Matches}} < 0$ then almost all the matching work will be performed manually. Ideally, $\text{Precision}_{\text{Matches}} = \text{Recall}_{\text{Matches}} = F\text{-Measure}_{\text{Matches}} = \text{Overall}_{\text{Matches}} = 1$.

5.2 *Results and Discussions*

Figure 2 displays the $\text{Precision}_{\text{Sets}}$, $\text{Recall}_{\text{Sets}}$, $\text{Overall}_{\text{Sets}}$ and $F\text{-Measure}_{\text{Sets}}$ obtained by SBP.

The findings (graph in Fig. 2) indicate that SBP reaches a high accuracy. The sets of words generated by SBP are very similar to the reference words sets. This is very promising.

The full matching results in terms of $\text{Precision}_{\text{Matches}}$, $\text{Recall}_{\text{Matches}}$, $\text{Overall}_{\text{Matches}}$ and $F\text{-Measure}_{\text{Matches}}$ are shown in Fig. 3.

The findings (graphs in Fig. 3) indicate quite similar results obtained by AML_{SBP}, $\text{PORSCHE}_{\text{SBP}}$ and $\text{COMA++}_{\text{SBP}}$ this is because SBP produces accurate and complete definitions for every schema element which helps improve the accuracy of the matches. The results also indicate that AML_{SBP}, $\text{PORSCHE}_{\text{SBP}}$ and $\text{COMA++}_{\text{SBP}}$ outperform AML, PORSCHE and COMA++, respectively, in terms of $\text{Precision}_{\text{Matches}}$, $\text{Recall}_{\text{Matches}}$, $F\text{-Measure}_{\text{Matches}}$ and $\text{Overall}_{\text{Matches}}$.

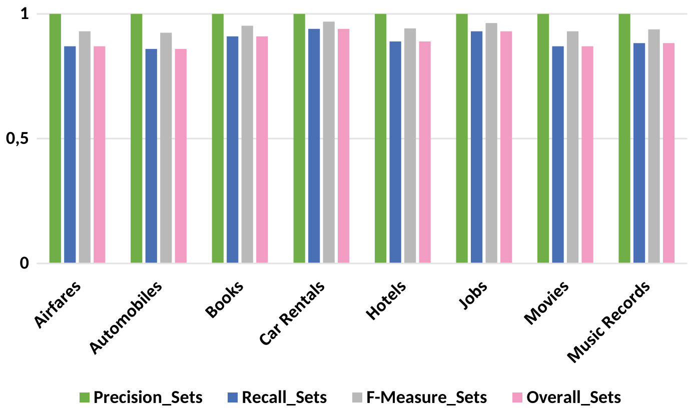


Fig. 2 The results obtained by SBP

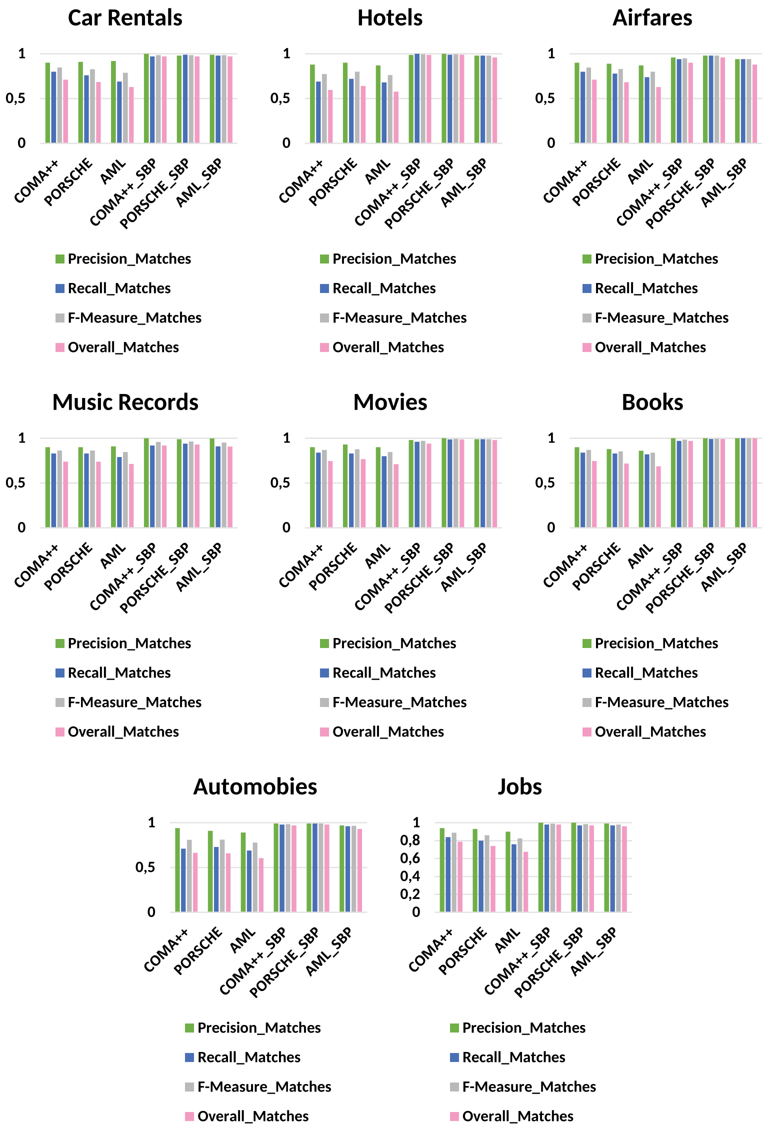


Fig. 3 The results obtained by AML_{SBP}, $\text{PORSCHE}_{\text{SBP}}$, $\text{COMA++}_{\text{SBP}}$, AML, PORSCHE and COMA++

6 Conclusions and Future Work

This paper has demonstrated that schema pre-matching is very critical for obtaining high matching accuracy. The state of the art matching tools do not take into consideration the fact that the sense of words often changes according to the context. Hence, this paper introduces an unprecedented a Sense-Based Pre-matching approach (SBP) that overcomes those limitations.

Given a schema S, our key idea is to prepare schema elements e for the actual matching. SBP captures the full meaning of e using both its vertical and horizontal contexts in S. It generates from e a words set θ that fully describes its meaning. This way, elements from different schemas are ready for the schema matching. This paper evaluated SBP on three state of the art matching tools over eight real-world domains. The results show that the matching tools applying SBP reach a superior matching accuracy. Future research work includes the following.

Consider other data structures. This paper focused mainly on hierarchical data structures and left other data structures for future work.

Study the impact of SBP on data source selection and ordering. Prior to answering the query, the system selects a subset of data sources that contain the complete or just a piece of the answer to the query (process is called source selection); next, the system orders the sources in a decreasing order of their coverage (given a query, a source coverage refers to the amount of answers contained in the source) (process is called source ordering). Thus, a future research direction would be to study the impact of SBP on source selection and ordering.

Work on schema matching. This paper focused on schema pre-matching; a future direction would be to come up with a holistic matching approach.

References

1. Zhang CJ, Chen L, Jagadish HV, Zhang M, Tong Y (2018) Reducing uncertainty of schema matching via crowdsourcing with accuracy rates. CoRR abs/1809.04017
2. Sutanta E, Wardoyo R, Mustofa K, Winarko E (2016) Survey: models and prototypes of schema matching. Int J Electr Comput Eng 2088-8708) 6(3)
3. Shvaiko P, Euzenat J (2005) A survey of schema-based matching approaches, pp 146–171
4. Yousfi, A, Yazidi MHE, Zellou A (2020) xmatcher: matching extensible markup language schemas using semantic-based techniques. Int J Adv Comput Sci Appl 11(8)
5. Bernstein PA, Madhavan J, Rahm E (2011) Generic schema matching, ten years later. PVLDB 4(11):695–701
6. Shvaiko P, Euzenat J (2013) Ontology matching: state of the art and future challenges. IEEE Trans Knowl Data Eng 25(1):158–176
7. Otero-Cerdeira L, Rodríguez-Martínez FJ, Gómez-Rodríguez A (2015) Ontology matching: a literature review. Expert Syst Appl 42(2):949–971
8. Ardjani F, Bouchiha D, Malki M (2015) Ontology-alignment techniques: survey and analysis. Int J Mod Educ Comput Sci 7(11):67

9. Faria D, Pesquita C, Balasubramani BS, Tervo T, Carriço D, Garrilha R, Couto EM, Cruz IF (2018) Results of AML participation in OAEI 2018. In: Proceedings of the 13th international workshop on ontology matching co-located with the 17th international semantic web conference, OM@ISWC 2018, Monterey, CA, USA, 8 Oct 2018, pp. 125–131
10. Saleem K, Bellahsene Z, Hunt E (2008) PORSCHE: performance oriented schema mediation. Inf Syst 33(7–8):637–657
11. Yazidi MHE, Zellou A, Idri A (2015) Fgav (fuzzy global as views). AIP Conf Proc 1644(1):236–243
12. Faria D, Pesquita C, Santos E, Palmonari M, Cruz IF, Couto FM (2013) The AgreementMaker light ontology matching system. On the move to meaningful internet systems: OTM 2013 conferences—confederated international conferences: CoopIS, DOA-trusted cloud, and ODBASE 2013, Graz, Austria, 9–13 Sept 2013. Proceedings, pp 527–541
13. Yousfi A, Yazidi MHE, Zellou A (2018) Assessing the performance of a new semantic similarity measure designed for schema matching for mediation systems. In: Computational collective intelligence—10th International conference, ICCCI 2018, Bristol, UK, 5–7 Sept 2018. Proceedings, Part I, pp 64–74
14. Yazidi MHE, Zellou A, Idri A (2013) FMAMS: fuzzy mapping approach for mediation systems. IJAEC 4(3):34–46
15. El Yazidi MH, Zellou A, Idri A (2012) Towards a fuzzy mapping for mediation systems. In: 2012 IEEE international conference on complex systems (ICCS), pp 1–4
16. Bourennani E, Bourque M (2019) A content-based schema matching tool. World Comput Sci Inf Technol J 9(5)
17. Ding G, Sun S, Wang G (2020) Schema matching based on SQL statements. Distrib Parallel Databases 38(1):193–226
18. Cruz IF, Antonelli FP, Stroe C (2009) AgreementMaker: efficient matching for large real-world schemas and ontologies. PVLDB 2(2):1586–1589
19. Do HH, Rahm E (2002) COMA—A system for flexible combination of schema matching approaches. In: Proceedings of 28th international conference on very large data bases, VLDB 2002, Hong Kong, 20–23 Aug 2002, pp 610–621
20. Massmann S, Engmann D, Rahm E (2006) COMA++: results for the ontology alignment contest OAEI 2006. In: Proceedings of the 1st international workshop on ontology matching (OM-2006) collocated with the 5th international semantic web conference (ISWC-2006), Athens, Georgia, USA, 5 Nov 2006
21. Miller GA (1995) Wordnet: a lexical database for english. Commun ACM 38(11):39–41
22. Aumueller D, Do HH, Massmann S, Rahm E (2005) Schema and ontology matching with COMA++. In: Proceedings of the ACM SIGMOD international conference on management of data, Baltimore, Maryland, USA, 14-16 June 2005, pp 906–908

Acceptance of Biometric Authentication Security Technology on Mobile Devices

William Ratjeana Malatji, Tranos Zuva, and Rene Van Eck

Abstract The development of mobile devices is quick and changes our daily personal and business lives. Every mobile user wants to be sure about individual data security, and for this reason, biometrics come into existence for mobile devices. Many studies were conducted on the acceptance of biometric authentication technology, but only a few of these studies focused on mobile devices-based biometry and the current study based on the mobile technology. To observe the reliability of the broadcast services, it is essential to offer better security for the biometry mobile phones. The limitations of this study were addressed by proposing a new mobile biometric technology acceptance model (MBTAM) that contains perceived humanness (PH), perceived interactivity (PI), and perceived social presence (PSP). The combined model for this quantitative study was tested on 302 mobile users through the distribution of the survey questionnaire, and examined by using the statistical package for social science (SPS). The results indicate that only one variable of the proposed model is not supported, which calls for further research. Furthermore, the functional elements of the research model become more prominent on the customer's intention to practice the mobile biometric device than the social elements. The research contributes to academic by suggesting new constructs that join together MBTAM to evaluate the possibility of mobile users to accept biometric authentication technology.

Keywords Mobile biometric technology acceptance model (MBTAM) · Perceived humanness (PH) · Perceived interactivity (PI) · Perceived social presence (PSP) · Statistical package for social sciences (SPSS)

W. R. Malatji (✉) · T. Zuva · R. Van Eck
Department of ICT, Vaal University of Technology, Andries Potgieter Blvd, Johannesburg 1911, South Africa
e-mail: villywr@gmail.com

T. Zuva
e-mail: tranosz@vut.ac.za

R. Van Eck
e-mail: rene@vut.ac.za

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_11

1 Introduction

In a technological era, mobile devices are most increasingly used for basic communications as well as a tool for managing individual issues and processing data obtained from anywhere at any time [1]. Over recent years, information access from mobile devices has become mainstream both in business and personal environments. The world is turning out to be more connected and every mobile user wants to be sure about individual data security [2].

Mobile device services assist as the base for business transactions but the traditional way of providing the security privileges is represented in terms of a mixture of alphanumeric and symbols. This ancestral process leads the users to avoid using mobile devices for reaching business data [3]. With the increase of its functionality including mobile banking, internet access, remote work, e-commerce, and entertainment, more confidential data is stored on these devices. For these reasons, biometrics comes in existence for mobiles [2].

To intensify the reliability of Wi-Fi services over mobile phones, a new trending and advanced technology have emerged that is biometric technology for mobile devices to promote the security levels [4]. Biometric technology refers to any technique that reliably uses measurable physiological or behavioral characteristics of distinguishing one individual from another [5].

Many studies were carried out on the acceptance of biometric devices and applications, users' attitudes towards such devices, and measurements of impact on performance. However, only a few of the studies focused on the factors that affect the acceptance of biometric devices [6]. Many studies have insisted on an investigation behind the biometric technology and stated the issues which are faced with user acceptance [7]. The acceptance of biometrics for other technologies still needs to be investigated deeply [8].

There were very few studies that measured the acceptance of biometric authentication technology on mobile devices. Therefore, this study efforts to regulate the reception of biometric corroborate technology on mobile devices.

The layout of the article is arranged in the following manner. Section 2 describes the related works of the proposed system. Section 3 denotes the significance of the study. Section 4 describes the methodology for the proposed system. Section 5 illustrates the results, and Sect. 6 reviews the discussions. Section 7 proposes the future scope of the research, and finally, Sect. 8 concludes the research work.

2 Literature Review

In literature, there have been many research studies on the acceptance of biometric devices and applications, users' attitudes towards such devices, and measurements of impact on performance. However, only a few of those studies focused on the factors that affect the acceptance of biometric devices [6]. Besides, each one of those examinations analyzed the adequacy of biometric procedures, however, do not contemplate the purposes for such acceptability. According to [9], many studies discussed the acceptance of technology, and the studies focused on technical issues such as algorithms, accuracy performance, etc.

The survey was carried out with 1206 respondents with the age of 18 years and above to find out the level of the acceptance of biometric technology (specifically facial recognition) from the Australian public [10]. This was achieved by asking how acceptable they thought it was if this technology was to be used in certain circumstances. It was found that 95% of respondents supported that the security can be used by airport staff as a way of passenger identification on police watch-lists. A similar report suggested with accuracy 92% of respondents have confirmed the security procedures chosen by the police for identifying the culprits in the criminal cases are of the video footage gathered through the security cameras. Among the survey report, quarter of the respondents weighed that this technology is not preferable for acceptance. One part of the respondents was bothering about the reflections of social media across these technologies (for example, Twitter, Facebook, and so on). It was found that 50% of the respondents declared this was an unacceptable technology to be applied [11].

Researchers conducted a review predicted on the physiological and behavioral biometric methods for user acceptance [12]. Later observed that these methods are rated very feeble in general except for fingerprint, voice, and hand geometry. All the above-mentioned studies have not been conducted based on mobile biometric devices.

According to [13], few studies have been conducted on mobile biometric devices and the good including the bad side of such devices were also discussed. Research conducted on both the pros and cons of the particular technique where there is no clear idea stated for the factors affecting the usage of biometric authentication technologies through mobile devices. The outcome of these factors is affecting the workplace, education, government sectors, and so on. Due to this report, there exists a phenomenon of technology for user acceptance [13].

Investigators studied modern mobile supporters towards their PDAs [14]. The particular biometric strategies were presented as elective confirmation measures to make sure about their mobile phones and observed that respondents reflected all techniques positively. The impediment of the investigation made by Clarke et al., Deane et al. and Furnell et al. [14–16] was that there was no attempt to comprehend the level of association concerning the members for biometrics on phones.

A portion of the effective determinants of biometric has been analyzed by Giesing [17] assessed the issues projected by the user and the social factors of biometric discovery. This examination leads to the new technology development towards the acceptance model designed by Davis [18].

3 Significance of This Study

By identifying the user acceptance issues from the research question, this research will at point consider how to address such issues to escalate the user acceptance of mobile biometric technology based on security. New devices are coming with biometric authentication security technology; however, few studies have tested the user acceptance of such technology on mobile devices. This examination will emphatically supplement the clients' consciousness of the biometric security reformation on cell phones. The findings of this study will assist decision-makers to be aware of the issues that affect users' decisions to welcome and utilize a specific system so that they would be capable of considering them during the development stage. It is hoped that this research would be beneficial to future researchers by providing them with helpful information about biometric authentication technology on mobile devices and some of their research questions may be answered by this study.

4 Methodology

4.1 Participants of the Study

Participants for this study were South African citizens in Vanderbijlpark. Three hundred and five (305) questionnaires were distributed to the target population. Only 302 responses were returned out of 305. The results of the demographic characteristics of the respondents are shown in Table 1.

Table 1 Questionnaire, source, and number of items

Constructs	Number of items	Source-citations
Perceived Ease of Use (PEOU)	4	Emily, Johnson and Carmen (2019)
Perceived Usefulness (PU)	4	Emily, Johnson and Carmen (2019)
Subjective Social Norm (SSN)	4	Barbara, Belanger and Schaupa (2017)
Perceived Humanness (PH)	3	Lankton, Knight and Tripp (2015)
Perceived Interactivity (PI)	3	Gao, Rau and Salvendy (2009)
Perceived Social Presence (PSP)	3	Lankton, Knight and Tripp (2015)
Intention to Use	2	Weng, Yang, Ho and (2019)
Actual Use of Mobile biometric device (AUMBD)	3	Asiimwe and Orebro (2015)
Trust	4	Cheng, Sun, Bilgihan and Okumus (2019)
Reliability	1	Tuunainen, Pitkanen and Hovi (2009)

4.2 Research Instruments

In this quantitative study, a simple random sampling technique was used to choose the participants. The items of this study in the survey questionnaire were constructed from the review of the related works that is appropriate to the research model. A five-point Likert—scale type measurement from one "strongly agree" to five "strongly disagree" was used in this study. After developing the questionnaire, it has been circulated to 30 participants (10% of the sample size) to ensure good clarity of questions, good length of instruments, and content completeness. The questionnaire is further sub-divided into two parts namely the first section and second section. The former part includes the details of the question linked to internet usage, technology expertise, demographics, and awareness of internet scams. The latter part consists of enquires about the estimation of the value of mobile biometrics (appropriate use of varied biometrics). Table 1 shows the questionnaire, source, and number of items. The questionnaire of this study was created based on the research framework derived from Ho et al. [19] shown in Fig. 1.

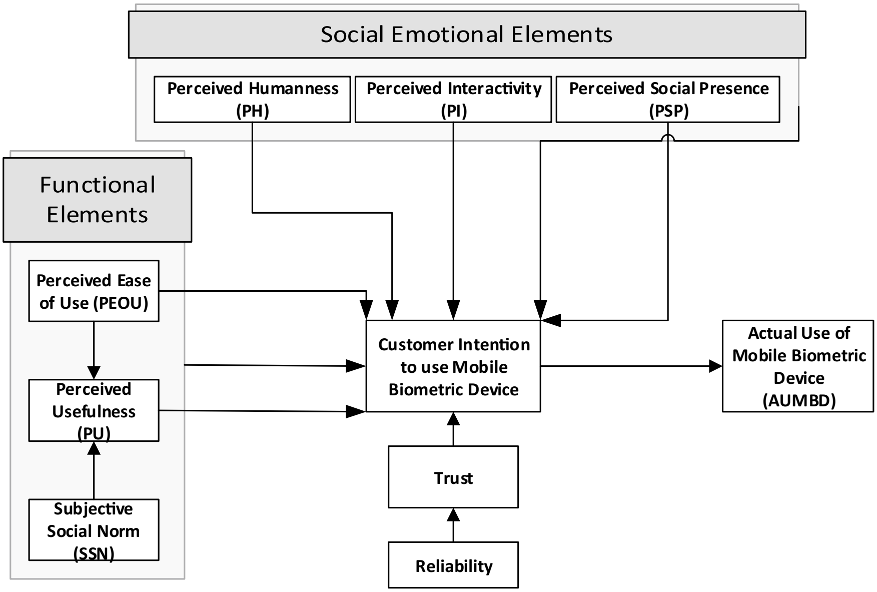

Social Emotional Elements
Perceived Humanness (PH)
Perceived Interactivity (PI)
Perceived Social Presence (PSP)
Functional Elements
Perceived Ease of Use (PEOU)
Perceived Usefulness (PU)
Subjective Social Norm (SSN)
Customer Intention to use Mobile Biometric Device
Actual Use of Mobile Biometric Device (AUMBD)
Trust
Reliability

Fig. 1 Proposed mobile biometric technology acceptance model. *Source* Ho et al. [19]

5 Results

5.1 Demographic Characteristics

The data that is presented in Table 2 of this study provides the demographic characteristics of the respondents on age, gender, race, employment status, and the level of the study. The results indicated that 186 respondents were male and 116 were female, which shows that the number of male respondents is larger when compared to the number of female respondents. The greater number of the respondents is between 20 and 30 years of age with 69.9%, while the smallest is between 51 and 60 years of age with 2.3%. Considering the nature of mobile devices, this imbalance is understandable, because most mobile users are usually the youth [20]. Of the different races that participated in the study, the results indicated that 198 respondents were black, 87 were white, and 17 were other races. Regarding the participant's employment status, the results show that 8. 6% were self-employed, 25.5% were employed, and 1.0% retired, while 62.3% were students, and 2.6% other. It was further indicated in the results that on the level of the study, the majority of the respondents were undergraduate students with 38.7%, and the lowest was primary with only 3% (Tables 3, 4, 5 and 6).

5.2 Statistical Analysis

The displayed research model in Fig. 1 was evaluated by employing the statistical package for social sciences. The primary solution for factor analysis of this study revealed that the model was appropriate for factor analysis. The assumptions were tested, and it was found that the data contained no outliers, and the level of close to normality was excellent. The produced results indicated that the dependent variables do not violate the presupposition of linearity. Moreover, the results indicated that there is no presence of homoscedasticity and there is no multicollinearity. This shows that the statistical inferences made regarding the data may be reliable. In this study, items reliability test was performed and it was found that the reliability analysis of all variables was fairly high, which showed that the internal consistency among variables was robust and greater. Furthermore, items validity test was performed and the results indicated the satisfactory level of the construct validity of items.

5.3 Regression Analysis

The objective of this work is to measure user acceptance of biometric authentication technology on mobile devices, the analysis will focus on the main variables of acceptance in our acceptance model. The key variables of the customer's for the purpose

Table 2 Respondents demographic informations

Variable	Frequency	Percent (%)
Gender		
Male	186	61.6
Female	116	38.4
Age		
19 and Below	14	4.6
20–30	211	69.9
31–40	53	17.5
41–50	17	5.6
51–60	7	2.3
61 and Above	0	0
Race		
Black	198	65.6
White	87	28.8
Other	17	5.6
Employment status		
Self-employed	26	8.6
Employed	72	25.5
Retired	3	1.0
A student	188	62.3
Other	8	2.6
Level of study		
Primary	1	0.3
Secondary	11	3.6
Undergraduate	117	38.7
Postgraduate	97	32.1
Other	76	25.2
Do you own a mobile device		
Yes	294	97.4
No	3	1.0
Owned it before	5	1.7
Have you used biometric authentication security before		
Yes	215	71.2
No	87	28.8
Would you prefer to use a mobile biometric device		
Yes	256	84.8
No	15	5.0

(continued)

Table 2 (continued)

Variable	Frequency	Percent (%)
Not sure	31	10.3
I have accessed the internet using a mobile biometric device before		
Yes	141	46.7
No	161	53.3
Total	**302**	**100**

Table 3 Regression results of PU, PEOU, SSN, trust, PH, PI, PSP and intention to use

Model		Unstandardized coefficients		Standardized coefficients	*t*	Sig.
		B	Std. Error	Beta		
1	(Constant)	0.026	0.166		0.154	0.877
	PU	0.350	0.066	0.324	5.278	0.000
	POEU	0.195	0.063	0.162	3.054	0.011
	SSN	−0.054	0.038	−0.070	−1.418	0.157
	Trust	0.350	0.066	0.311	5.103	0.000
	PH	0.132	0.060	0.126	2.196	0.029
	PI	0.196	0.064	0.166	3.070	0.002
	PSP	0.211	0.052	0.229	4.030	0.000

[a]Dependent variable: intention to use

Table 4 Regression results of intention to use and AUMBD

Model		Unstandardized coefficients		Standardized coefficients	*t*	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.114	0.141		14.988	0.000
	Intention to use	0.224	0.072	0.177	3.107	0.002

[a]Dependent variable: actual use

Table 5 Regression results of PEOU, SSN, and PU

Model		Unstandardized coefficients		Standardized coefficients	*t*	Sig.
		B	Std. Error	Beta		
1	(Constant)	0.427	0.132		3.228	0.001
	POEU	0.589	0.048	0.576	12.340	0.000
	SSN	0.128	0.032	0.180	3.961	0.000

[a]Dependent variable: PU

Table 6 Regression results of reliability and trust

Model		Unstandardized coefficients		Standardized coefficients	*t*	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.207	0.182		6.634	0.000
	Reliability	0.466	0.052	0.468	8.971	0.000

[a]Dependent variable: trust

of using the mobile biometric devices (Intention to use) are PEOU ($\beta = 0.162$; $p < 0.05$), PU ($\beta = 0.324$; $p < 0.01$), PH ($\beta = 0.126$, $p < 0.05$), PI ($\beta = 0.166$; $p < 0.05$), PSP ($\beta = 0.229$; $p < 0.01$) and trust ($\beta = 0.311$; $p < 0.01$). The results indicate that trust and PU are the most important variables in explaining customer's intention to utilize the mobile biometric devices (Intention to use). Intention to use on its own is a key variable to AUMBD with ($\beta = 0.177$; $p < 0.05$). It is indicated in the results that PEOU is the most important variable that explains PU ($\beta = 0.576$; $p < 0.01$) succeeded by SSN ($\beta = 0.180$; $p < 0.01$). Moreover, reliability is the most important variable that explains trust with ($\beta = 0.468$; $p < 0.01$). The sum of functional elements of our model indicates that PEOU, PU, and SSN altogether, strongly explain intention to use with ($\beta = 0.390$; $p < 0.01$) (Fig. 2).

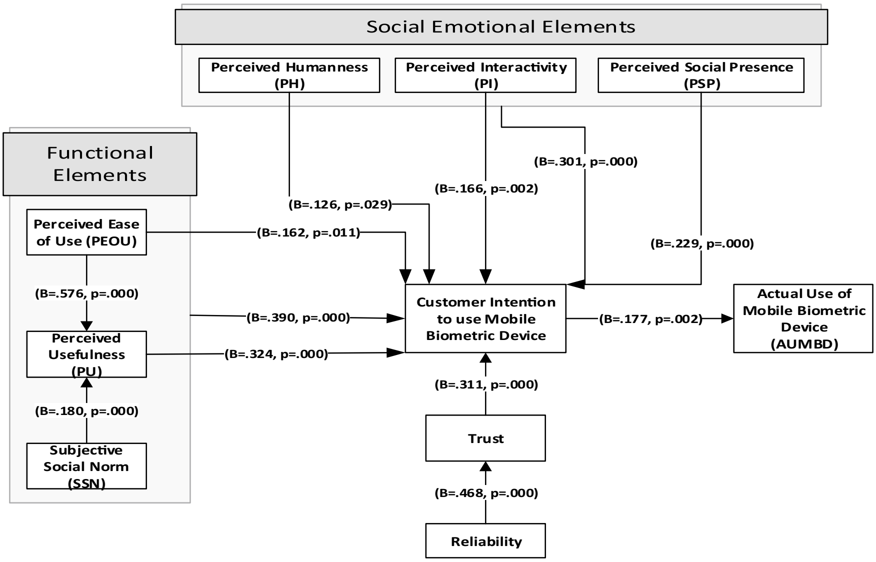

Social Emotional Elements
Perceived Humanness (PH)
Perceived Interactivity (PI)
Perceived Social Presence (PSP)
Functional Elements
Perceived Ease of Use (PEOU)
Perceived Usefulness (PU)
Subjective Social Norm (SSN)
(B=.301, p=.000)
(B=.166, p=.002)
(B=.126, p=.029)
(B=.162, p=.011)
(B=.229, p=.000)
(B=.576, p=.000)
(B=.390, p=.000)
(B=.324, p=.000)
(B=.180, p=.000)
Customer Intention to use Mobile Biometric Device
(B=.177, p=.002)
Actual Use of Mobile Biometric Device (AUMBD)
(B=.311, p=.000)
Trust
(B=.468, p=.000)
Reliability

Fig. 2 Proposed Model for this study

6 Discussion

The overall mobile biometric acceptance model that is proposed in this study is validated. Starting with the functional elements (PEOU, PU, and SSN) of the model, the results indicated that PEOU has a positive influence on customers' intention to use mobile biometric devices (intention to use), and these results were supported by Suki and Suki [21]. This is an indication that when PEOU increases also intend to use increases. The results show that PU obtained impacts the positive plan to accept the usage of mobile, and these results are also in line with [21]. However, SSN on its own was not supported in this study. These same results were found on Chao [22]'s study on "factors determining the behavioral intention to use mobile learning: an application and extension of the ATAUT model." Based on the obtained results, it is concluded that PEOU and PU can be kept and used in future research to measure the acceptance of biometric authentication security technology on mobile devices. Although SSN is not supported, the variable on its own influence PU, moreover, the sum of all functional elements indicates a very strong influence on intention to use. Therefore, the conclusion cannot yet be made on whether the variable must be removed or not.

The social elements (PH, PI, and PSP) of the proposed model are all supported. It was indicated by the results that PI, PH, and PSP have a positive intention to use, and the results of these three variables are supported by Lankton [23]. Therefore, it is concluded that these variables can be kept and used in future research to estimate the user acceptance of biometric authentication technology on mobile devices [22]. Trust on its own is strongly influenced by reliability. Reliability is the most important variable that explains trust, and these results are in line with [7]. Intention to use on its own has a positive influence on AUMBD which is supported by Suki and Suki [21]. Based on these results, the conclusion can be made that trust, reliability, intention to use, and AUMBD can be kept and used in future research to measure the acceptance of biometric authentication technology on mobile devices.

7 Limitations and Suggestions for Further Research

This study focused on the two limitations as follows. Firstly, the study focused on the acceptance of biometric authentication technology on mobile devices only. Further research must be carried out on the acceptance of biometric authentication on other existing technologies except for mobile devices. The second important limitation of this study concerns gender and age of the respondents. The majority of the respondents for this study were male, and the highest age group of respondents was between 20 and 30. This brings about an issue of unbalanced results. Generally, both males and females in different age groups nowadays are using mobile devices. The conceptual framework used in this study should also be tested on the acceptance of biometric authentication technology on other existing technologies.

8 Conclusion

This study aimed to measure the acceptance of biometric authentication technology on mobile devices. The model that was used in this study proved to be valid, suitable, and supported. The researcher suggested that further research must be done especially using the variables that were supported in the model. The results and findings of this research showed that the majority of respondents acknowledged or are willing to accept biometric authentication technology to be used as security on mobile devices. However, further research needs to be conducted in this area.

Acknowledgements The author like to appreciate Professor Tranos Zuva and Doctor Rene Van Eck from the Vaal University of Technology for their support and supervision in this research. This is the revised version of the earlier paper. The author also likes to thank the IMITECH-2020's two anonymous reviewers for their valuable suggestions and comments throughout the research.

References

1. Wang H, Liu J (2009) Mobile phone-based health care technology. Recent Patents Biomed Eng 2(1):15–21
2. Kadena E, Ruiz L (2018) Adoption of biometrics in mobile devices. Obuda University, Doctoral school on safety and security sciences, Budapest Hungary. technologies to support teachers and improve practice
3. Bao P, Pierce J, Whittaker S, Zhai S (2011) Smartphone use by non-mobile business users. In: MobileHCI, Stockholm, Sweden. Attitudes and Practices: Computers & Security, vol 24, no 7, pp 519–527
4. Clarke N, Furnell S (2005) Authentication of users on mobile telephones
5. Kaur G, Kaur D (2013) Multimodal biometrics at feature level fusion using texture features. Int J Biometr Bioinform 7(1):58–73
6. James T, Pirim T, Boswell K, Reithel B, Barkhi R (2017) Determining the ıntention to use biometric devices: an application and extension of the technology acceptance model. J Organ End User Comput 18(3)
7. Chau A, Jamieson R (2004) Biometrics acceptance-perception of use of biometrics. Assoc Inform Syst
8. Uzoka FE, Ndzinge T (2009) Empirical analysis of biometric technology adoption and acceptance in Botswana. J Syst Softw 82:1550–1564
9. Chau A, Jamison R (2004) Biometric acceptance-perception of use of biometrics. In: ACIS 2004, Proceedings
10. Newspoll (2012) Rite aid deployed facial system in hundreds of Australia public. J Organ End User Comput 4(10):110–115
11. Unisys.Unisys security index report australia: facial recognition.https://www.unisyssecurityindex.com/system/resources/uploads/101/original/Australian2012.pdf
12. Miltgen L, Popovic C, Oliveira T (2013) Determinants of end-user acceptance of biometrics: ıntegrating the Big 3 of technology acceptance with privacy context. Decis Support Syst 56:103–114
13. Vrana R (2018) Acceptance of mobile technologies and m-learning in higher education learning: an explorative study at the faculty of humanities and social science at the University of Zagreb. Department for İnformation and Communication Science

14. Clarke NL, Furnell S, Rodwell PM, Reynolds PL (2002) Acceptance of authentication methods for mobile telephony devices 21(3):220–228
15. Deane F, Barrelle K, Henderson R, Mahar D (1995) Perceived acceptability of biometric security systems. Comput Secur 14(3):225–231
16. Furnell SM, Dowland PS, Illingworth HM, Reynolds PL (2000) Authentication and supervision: a survey of user attitudes. Comput Secur 19(6):529–539
17. Giesing I (2020) User perceptions related to ıdentification through biometrics within electronic business. University of Pretoria. https://upetd.up.ac.za/thesis/available/etd-01092004-141637/. Accessed 17 Feb 2020
18. Davis FD (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quart 13:319–339
19. Ho G, Stephens G, Jamieson R (2003) Biometric authentication adoption ıssues. In: Presented at the proceedings of the 14th Australasian conference on ınformation systems, Perth, Western Australia, 26–28th November 2003
20. Hosokawa R, Katsura T (2018) Association between mobile technology use and child adjustment in early elementary school age: Plos One J 13(7)
21. Suki NM, Suki NM (2011) Exploring the relationship between perceived usefulness, perceived ease of use, perceived enjoyment, attitude and subscribers' intention towards using 3G mobile services. J Inf Technol Manage (2011)
22. Chao C (2019) Factors determining the behavioural intention to use mobile learning: an application and extension of the UTAUT model. Front Psychol 10
23. Lankton M, McKnight DH, Tripp J (2015) Technology, humanness and trust: rethinking trust in technology. J Assoc İnform Syst 16(10) (2015)

Impact of Agile Methodology Use on Project Success in South African Banking Sector

Makoena Moloto, Anneke Harmse, and Tranos Zuva

Abstract Agile methodology has become the most used software development methodology in different software communities. Besides the extensive benefits that agile methods offer, it also posseses several critical issues. Most studies have reported that, when comparing structured methodologies with agile methodologies, agile improves software quality, whereas other studies have contradicted and questioned the claimed benefits. In contrast, others argued that it does not have enough empirical evidence. This study aims at investigating the impact that agile method implementation on project success in the South African banking sector. The survey was conducted through questionnaires. The theoretical model for analyzing the impacts of agile method implementation on project success was proposed. Reliability was tested using Cronbach's alpha. The data were quantitatively analyzed by using correlation and regression approaches that use statistical package for social sciences. The results indicated that agile methodology implementation positively influences the project success. Furthermore, this study has indicated that reduced upfront planning, iterative delivery, environmental feedback and technical feedback as individual agile method use components positively impact project success in the South African banking sector.

Keywords Agile methodologies · Agile impact · Agile benefits · Project quality · Project success · Project performance

1 Introduction

The success of the project is usually measured by the perceived benefits to the user and the significant positive impact the project brings to the organization together with the return of investment [33]. Lately, industry experts have been keen to enhance performance at both the organization level and project level by adopting agile practices because agile enables organizations to stay more flexible [36]. Organizations

---

M. Moloto (✉) · A. Harmse · T. Zuva
Faculty of Applied and Computer Sciences, Vaal University of Technology, Vanderbjilpark, South Africa
e-mail: koenasm@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_12

following a plan-driven approach are primarily unable to accommodate changes during software development; they face poor communication between customers and developers, poor project management and high project cost [16]. They are seeking ways to stay competitive in the market share, to comply with relevant standards and deliver software on time and therefore adopt and use agile methodologies in their projects [7]. Agile seems to be the rational approach to deal with the instability of the market and to meet dynamic business needs [36].

Software development industries around the globe started changing their development process to agile since its manifesto was introduced over a decade ago [16]. Agile has formed in a manner that software development communities put people first to ensure transparency, trust and teamwork between stakeholders and to solve problems that the traditional software development methodologies could not solve [35]. It was created to ensure projects do not fail anymore because of people's problems, as demonstrated by most studies [6]. People are creative and that is why agile depends on them over a process and organizational maturity [6].

Agile methodology has been adopted in many software development projects due to its ability to better cope with frequent changes in requirements maturity [8]. It can improve the management of the development process and developer or user relationships [5]. Teams move to agile to improve quality, to reduce defects rates, to comply with standards, to deliver systems in time, under budget and to have a better interaction with customers [7]. Asnawi et al. [3] indicated that even if the use of agile is still emerging in Malaysia and its use is still forming, benefits are showing in the adoption of the methods. Agile ensures the delivery of quality software products on time [15].

1.1 Problem Statement

Regardless of how widely agile methodologies have been adopted and used as a means of solving the issues plan-driven methodologies could not solve, with a number of authors pointing the benefits and advantages it brings, very less empirical studies have been done to support the argument on project success [24] Many authors have written about the benefits agile brings to software engineering, many contradicted and questioned the claimed benefits. In contrast, others argued there is not enough empirical evidence of the claimed benefits [25].

Agile has been significantly adopted, progressed and used in the information system industries. Still, there are no neutral academic studies to its advantages. All the benefits claimed need more empirical studies [17]. The claims that are made by agile communities do not have enough scientific evidence supporting them [20]. Owen and Koskela [24] stated that because projects tend to be complicated by nature, it is essential to verify that when comparing agile to other methodologies, it improves projects for better and ensure the claimed benefits are true.

It is for these reasons that this study aimed to evaluate the impact of agile methodology use on project success in the South African banking sector.

Research questions. How does agile methodology use impact project success in the South African banking sector?

Aims and objectives. This study aimed to measure the impact of agile methodologies uses on project success in the South African banking sector.

The layout of this paper is as follows:

1. Introduction; to give an introduction to our research.
2. Related work; finding out what research has been done and if there is enough evidence on the impacts of agile methodology use on project success.
3. Proposed work; broadly outlining the research plan, research procedure, the research design and methods used in this paper.
4. Results analysis; explaining the findings of the research in detail and finally, the conclusion summarizing the research study.

2 Related Work

In the last years, agile development has been the most adopted process that has gained strong momentum [15]. A need for quality, efficient, reliable and useable software is continuously growing across the globe. As a result, a need for geographically accommodative software development methodologies rises [30]. Waterfall methodology exposes significant project risk because it relies on extensive and comprehensive planning; it also does not favor customers in changing their needs due to its sequential software development life cycle [9]. Fast software delivery, stable development process, ambiguous and/or evolving requirements are the popular motivations for adopting agile [37].

Agile is aiming at implementing the projects responsive to change by emphasizing less on the development of the rigorous project requirements and plans, but more on revising the requirements after every cycle [19].

The idea that agile is undisciplined and risky is a myth. Agile principles like stand up meetings keep communication and everyone engaged and informed of what is happening [28]. Daily stand-ups meetings improve communication between the team members, and this leads to trust between them and therefore improves the return of investment performance. They share knowledge and become self-organized [10]. It practices better and efficient communication methods between developers and customers [14]. When using agile, communication strategies improve, and stakeholders become more engaging. Hence, issues and faults get addressed in good times, and costs are reduced [2].

Agile iterations help communicate the progress of the project quickly and transparently than traditional methodologies. Software changes are then picked up as soon as possible because customers have a close relationship with the team members and progress feedback is given after every iteration [33]. To build high-quality projects on time, agile encourages customers to meet with developers regularly to validate and verify the requirements [29]. Velumani [36] agrees that agile methodologies helped many organizations to deal with the volatility of the market and to meet dynamic business needs after employing it in the past decade. When [15] compared agile techniques with traditional approaches, 89.6% of participants stated that with agile productivity was higher, 84.8% of participants stated that the quality is higher, and 86.5% stakeholders were highly satisfied. Agile helps teams become disciplined. They become better in communication, how to set goals and how to define acceptance criteria. Coordination and prioritization of activities can be enhanced through an integrated environment where testing can occur anytime, and system architecture should make parts of the system as independent as possible to avoid code conflicts across all project teams. Project organization structure hierarchy may also be introduced to enhance coordination [19].

People who successfully follow agile processes acquire their benefits such as delivery time improvements, quality improvements and development costs reduction [22]. Sidky and Arthur [29] suggested that it is important for organizations to analyze their past projects and define success factors. They need to ensure their organizational culture fits well with agile. Sidky and Arthur [29] further emphasize that people must know and understand agile principles through training and swallow it into their work culture. Vijayasarathy and Turk [37] surveyed early adopters in agile software development and mentioned that organizations that implemented agile methodologies had given positive feedback. Increased productivity with less defects, maintainable and extensible code with high test coverage and reduced time and costs are amongst the positive feedback given by the organizations which participated in a survey [37]. On a personal level, they mentioned that morale has improved with better collaboration and their customers are very satisfied, which is the key benefit of using agile development methodology. Agile adoption should not only be on an organization level but also a project level and be implemented within teams, not at a different time [28].

Agile software is deployed continuously, and the teams get to test the system regularly and continuously as they test every phase and therefore detect the issues at an early stage before they become more severe. Fewer defects mean better quality, which means more customer satisfaction [2]. Unlike traditional methods, faults are detected early and not at the end of the whole software development emphasized [11]. High-quality software is built because of the trust developers have with customers who also improves the initial states of the methodology. The owner's needs are met through onsite interaction from customers. Costs are reduced because changes are implemented early and the software is delivered incrementally [2].

Different projects in software development industries are developed and prototyped using agile methodology due to its ability to satisfy the changes in requirements, unlike the traditional development methods, which are not efficient in that regards [10]. Agile sees coding as the core of software development; it has, however, been criticized that code alone can lead to the loss of information due to lack of detailed documentation and models, especially in large complex systems. Those in favor argue that code is the base and the pillar, the only deliverable that matters in software creation and evolution, rather than analysis, designs and documentation [34]. Unlike traditional methods which spare a huge amount of time planning and documenting, and additional great effort of rework, defects are usually discovered early while they are embedded because products are verified during creation, which then becomes less costly, less time consuming and less effort to identify and remove them [7]. Plan-driven methodologies waste time documenting bulk requirements that might have to be reworked as the customers change their needs [9].

Generally, agile or scrum provides better results in software development. In the long term, it also increases quality and productivity [15]. When [35] made his research on The Adoption of Agile Software Development Methodologies by Organizations in South Africa, he mentioned that project visibility improved, team management regarding changing priorities improved, market time was faster, software quality and team morale also improved. Organizations switching from the traditional waterfall process model can increase the morale of the team and decrease large-scale software application risks [9]. An open-minded team, together with a strong scrum master, also helps to have a successful agile software development project [15].

When [22] researched adopting agile practices when developing medical device software, they mentioned that the project was completed seven per cent faster than if it were to be implemented using plan-driven approaches and stakeholders confirmed it would have overrun by 14% with budget overrunning by approximately seven per cent. A number of the benefits, including but not limited to productivity, project visibility and software quality, have been reported on adoption and use of agile methodologies [17]. The plan-driven methodologies processes do not allow changes to be adopted during implementation, therefore, end up with possibilities of obsolete product implemented, with costs increases and wasted resources because the initial requirements are no longer desired [9].

Agile teaches people to do just enough documentation needed. For people who have been on traditional methods for a while, where comprehensive and detailed documentation is the key, through training and educations, they will have a smooth, agile adoption [18]. Incorporating agile in software development can benefit organizations [9]. “Many of us have encountered people and teams who claim to be ‘agile,’ but who are just making this claim to avoid planning, documentation and design. If agile methods were simply about avoiding those aspects of software and systems delivery, then agile methods would be entirely inappropriate for safety–critical systems development. That is a misrepresentation of the intent and the practice of agile methods” [7]. Agile results in a high project satisfaction because it is adaptive and open to continuous customer feedback while accommodating changes due to its nature to have repeated requirements gathering, production deployable working source code, implementation of functionality and system testing [9].

Because of the freedom given by most agile methods, most people do follow the manifesto but not adhere to the specific agile methodology [15]. When using agile, projects become efficiently delivered because they are delivered in short iterations [13].

2.1 Agile Project Success

Project cost estimations using traditional software development methodologies have always been difficult. However, delivering the project on time, under budget, within scope has always been the critical aspect of a successful project [13]. Serrador and Pinto [27] used two dimensions to define project success, Project efficiency: defined by whether the scope goals of the project are met and whether the project is delivered on time within budget and Stakeholder success: defined as the best judges of the overall success measuring whether the project expectations are met and therefore satisfying the stakeholders. Project managers always find making software successful and increasing the customer satisfaction challenging, especially in an environment where changes come first not taking consideration of scope, time, cost and quality and not thinking of the positive or negative impact they might have to the project [38].

For a project to be declared successful, the scope has to be delivered on time under budget and be of good quality [23]. Projects that are not delivered on time may have sales lost and higher development costs, which may also result in customers not being satisfied emphasized [32]. It is crucial for organizations to understand the critical success factors of the projects to improve project management strategies, cost benefits, profitability and productivity. By so doing, they will positively impact social change and benefits management, employees and customers [23]. Planning and managing scope, time, cost, quality, risk, human resources, stakeholders and procurement management are the primary keys of project management with time and cost being the main issues [38].

Because systems are delivered over budget and costly, methodologies that can manage system development are needed [2]. The inappropriate choice of a project management methodology may contribute to the project being unsuccessful [1]. Agile helps the project manager and business development manager become more prepared for the ever-changing requirements of the clients. It also equips them with the understating of scope, cost and budget estimates, to avoid using the patterns without an understanding [13]. Requirements, scope and timeline change profoundly affect cost estimation in most cases. Agile companies, unlike companies using traditional methods, have their cost estimation at the beginning of every iteration, which helps them have accurate estimates easily. For the project to succeed, an organization must have agile engineering practices continuously integrated to deliver quality scope under budget, finalize all requirements before starting development to avoid bugs and glitches, ensure there are automated tests and builds in place, engage and motivate the team to achieve higher quality project deliverables, have stronger technical practices combined and have time-boxed iterations [28]**.** Managers must avoid having failing projects due to the methodologies which do not accommodate changes and always try and maximize profits and minimizes costs [2].

Regardless of software development difficulties in estimating costs, projects must be delivered under budget, within time, and the requirements must be met [13]. When using agile, projects become efficiently delivered because they are delivered in short iterations [13]. Agile put people first over the process. Creative and innovative team members help project success [2].

3 Proposed Work

The quantitative research method is used in this study. It is considered conclusive in its purpose, and it can be used to measure statistical results that are interpreted objectively [21]. This study used inferential statistic to analyze the impact of agile methodology use on project success in the South African banking sector. A research model in Fig. 2 for effectively analyzing the impacts of agile method use on project success in the South African banking sector was proposed and its effectiveness measured. Our model was derived from the theoretical model for analyzing the impacts of agile method adoption on project success by Tripp [33]. As seen in Fig. 2. The Model consists of two main variables, agile method use and project success. Agile method use has multiple components, and they are discussed below:

Reduced upfront planning determines how much time spent is reduced by the team before beginning the work. The reduced upfront planning is to directly impact the project performance by reducing time to initial feedback and waste from planning tasks too far in advance [33]. Iterative delivery determines how much time has taken the team to deliver the functional work in an iteration. It reduces ambiguity in requirements due to the feedback received from the users, which is from the working system rather than the abstract documents [33]. Environmental feedback determines the mechanisms utilized by the team to obtain feedback from the customers and stakeholders [33]. Technical feedback determines to what level the mechanisms have been utilized by the team to ensure that the system is functioning properly [33]. Project success is dependent on agile method use, and its components are discussed below.

Project performance determines whether the project results have matched the defined goals of scope schedule and budget. It is measured using the project budget outcome, project time outcome scale and project scope outcome [33]. Product quality determines whether the system is useful, reliable, complete, effective, suitable and accurate as perceived [33]. Project impacts determine to what extent or how positively has the project impacted the organization as perceived. The perceived benefits of the system are measured with two dimensions: Customer satisfaction and organizational impacts [33].

3.1 Sample Size and Profile Participants

The criteria for participating in our study were that a team is using agile software development methodology operating in the application domains of a South African banking sector and it had delivered software to a customer at least once. The respondents varied from a small team with less than 10 team members to large teams with several hundred team members. One hundred and fifty responses were collected, and 41 were incomplete and discarded. The remaining 109 complete responses were used for analysis. Our participants comprised mainly but not exclusively of IT stakeholders, IT management and team members (software developers, quality assurance personnel, project managers, business analysts, system analysts and architects).

3.2 Data Collection

The questionnaire was the key data collection instrument in this study. Data were quantitatively collected in South African banks. The data collection was administered using primary data. A non-probability purposive sampling method has been used for data collection. This sampling method was used because not all project teams in the South African banking sector are using the agile method. It was appropriate to make use of a purposive sample in this survey because sampling from outside the population, which is not using the agile development method, would be inappropriate. For us to gather as much data as possible, Snowball sampling is also used and asked the participants to indicate any other potential participants who are working in the banking sector on different projects and are using an agile methodology. For our respondents to take part in the survey, the questionnaire is placed online using google forms.

3.3 Data Analysis

The statistical package is used for Social Sciences (SPSS) V25 software to analyze our data. Figure 1 shows the methods used to evaluate our data. As shown in Fig. 1, the common weaknesses of the survey study (validity and reliability) were mitigated by using Cronbach's alpha coefficient and Kaiser–Meyer–Olkin (KMO). According to [12], Cronbach's alpha is not a statistical test. It measures the relationship between many items in a group, how closely related they are and therefore measure consistency. Cronbach's alpha coefficient ranges between 1 and 0. The closer the value approaches 1, the higher the reliability, and the closer it approaches 0, the lower the reliability [31]. George and Mallery [12] provided the following Cronbach's alpha rules of thumb: _ > 0.9—Excellent, _ > 0.8—Good, _ > 0.7—Acceptable, _ > 0.6—Questionable, _ > 0.5—Poor and _< 0.5—Unacceptable. To measure how adequate our sampling is, KMO is used. "KMO test reflects the sum of partial correlations relative to the sum of correlations" [4]. A measure of >0.9 is marvelous, >0.8 is meritorious, >0.7 is middling, >0.6 is mediocre, >0.5 is miserable, and <0.5 is unacceptable [12].

To test how strong the relationship is between our variables, correlation analysis has been used. Correlation analysis tests how strong the relationship is between two continuous variables given as a coefficient between −1 and 1. A score ranging −1

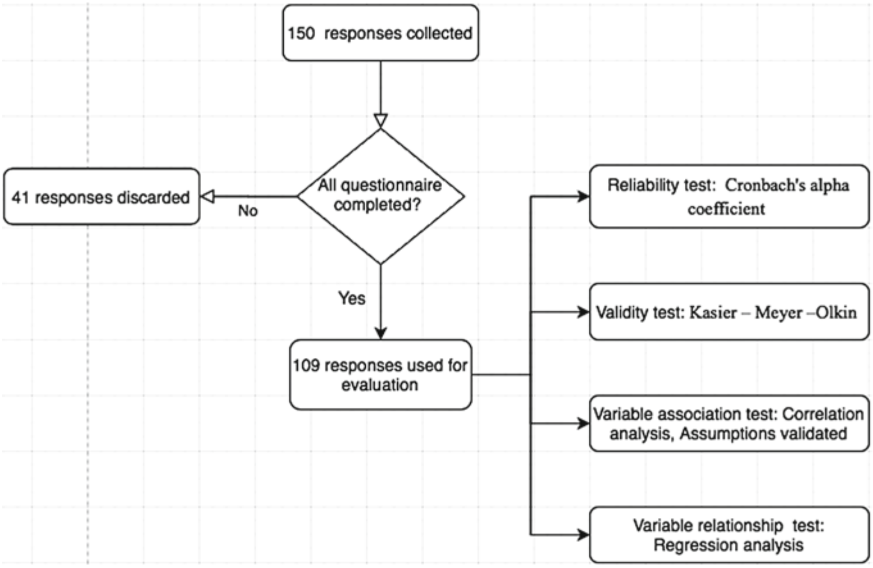

150 responses collected
All questionnaire completed?
No
41 responses discarded
Yes
109 responses used for evaluation
Reliability test: Cronbach's alpha coefficient
Validity test: Kasier – Meyer –Olkin
Variable association test: Correlation analysis, Assumptions validated
Variable relationship test: Regression analysis

Fig. 1 Evaluation method

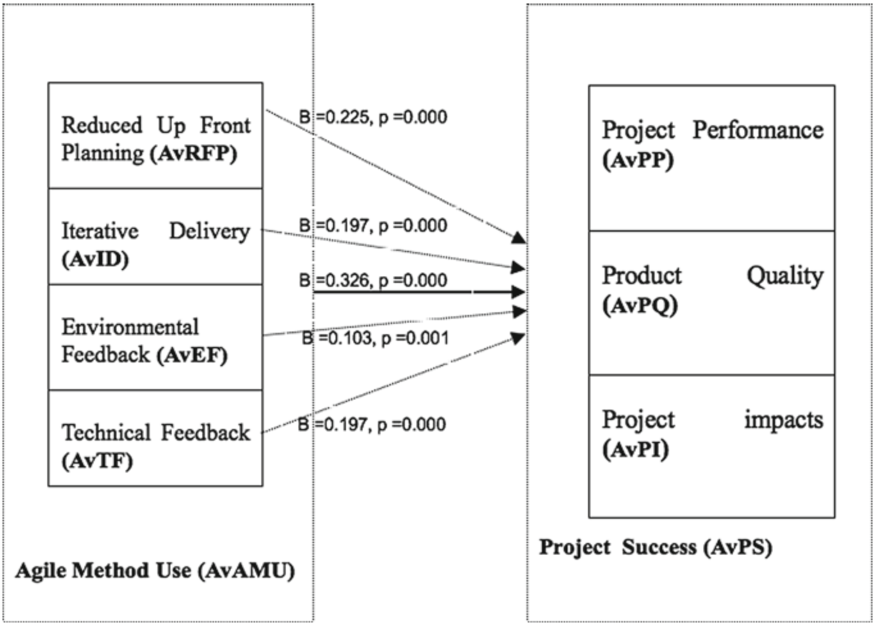


Fig. 2 Theoretical model for analyzing the impacts of agile method adoption on project success [33]

1 is perfect, −0.9 to −0.7 0.7 to 0.9 is strong, −0.6 to −0.4 0.4 to 0.6 moderate and −0.3 to −0.1 0.1 to 0.3 weak [12].

Regression analysis is also used to indicate how significant the relationship is between our independent and dependent variables. Regression indicates the effects of the relative strength of different independent on a dependent variable [26].

4 Results Analysis

The reliability and validity were measured, and the following results were recorded as seen in Tables 1 and 2.

A score 0.801 is indicated in our results. As [12] indicated, our results show high reliability of our variables.

KMO was used to test the validity of our study and indicated a value sampling adequacy of 0.687. According to [12], this is acceptable. These variables have proven to be statistically significant at the 1% level.

4.1 Correlation

To test how strong the relationship is between our variables, a correlation analysis has been performed. The results are shown in Table 3.

From Table 3, it can be seen that there is a positive relationship between the individual agile method use variables and project success. The correlation showed that reduced upfront planning (Pearson correlation 0.467, Sig. 0.000), iterative delivery (Pearson correlation 0.356, Sig. 0.000), environmental feedback (Pearson correlation 0.327, Sig. 0.000), technical feedback (Pearson correlation 0.403, Sig. 0.000), agile

Table 1 Cronbach's alpha

Reliability statistics	
Cronbach's alpha	No. of items
0.801	9

Table 2 Kaiser–Meyer–Olkin

KMO and Bartlett's test		
Kaiser–Meyer–Olkin measure of sampling adequacy		0.687
Bartlett's test of sphericity	Approx. Chi-Square	2728.856
	Df	1225
	Sig.	0.000

Table 3 Correlation analysis for agile method use variables and project success

Correlations						
		AvRFP	AvID	AvEF	AvTF	AvPS
AVPS	Pearson correlation	0.467**	0.356**	0.327**	0.403**	1
	Sig. (2-tailed)	0.000	0.000	0.001	0.000	
	N	109	109	109	109	109

**Correlation is significant at the 0.01 level (2-tailed)

method use (Pearson correlation 0.533, Sig. 0.000) have a positive relationship with project success.

4.2 Regression

The direct effects of the agile method use on project success were tested. Regression analysis was used to determine the actual contribution of agile method use on project success, and the results are shown in Tables 4 and 5.

In Table 4, it is indicated that the influence of the individual agile method use variables on project success is positive. The following variables were noted; reduced upfront planning influences project success with a B coefficient values of 0.225 and *p*-value of 0.000, iterative delivery influences project success with B coefficient value of 0.197 and *p*-value of 0.000, environmental feedback influences project success with a B coefficient value of 0.103 and *p*-value of 0.001, and technical feedback influences project success with a B coefficient value of 0.197 and *p*-value of 0.000. These variables have proven to be statistically significant at the 1% level. The equation for the use of agile method variables to project success is as follows:

$$AvPS = 0.225AvRFP + 0.197AvID + 0.103AvEF + 0.197AvTF + 2.5, \tag{1}$$

where AvPS is the project success depends on individual agile method use components. AvRFP is reduced upfront planning determining the time reduced by the team, AvID is iterative delivery determining the time taken by the team to deliver a functional code, AvEF is environmental feedback determining the feedback mechanism used by the team, and AvTF is technical feedback determining the feedback provided by the team mediating technology.

Using regression analysis, the direct effects and the actual contribution of agile method use as a whole on project success were tested. The results are seen in Table 5.

The results indicated that the influence of the extent of agile method use as a whole on project success is positive. Table 5 indicates that the agile method use positively

Table 4 AVPS dependent on AvAMU components (AvRFP, AvID, AvEF, AvTF)

Coefficients[a]								
Model		Unstandardized coefficients		Standardized coefficients	t	Sig.	95.0% confidence interval for B	
		B	Std. error	Beta			Lower bound	Upper bound
1	(Constant)	2.423	0.117		20.796	0.000	2.192	2.654
	AvRFP	0.225	0.041	0.467	5.461	0.000	0.143	0.306
	AvID	0.197	0.050	0.356	3.940	0.000	0.098	0.296
	AvEF	0.103	0.029	0.327	3.584	0.001	0.046	0.159
	AvTF	0.197	0.043	0.403	4.561	0.000	0.111	0.283

[a] Dependent variable: AVPS

Table 5 AVPS dependent on AvAMU

Coefficients[a]

Model		Unstandardized coefficients		Standardized coefficients	t	Sig.	95.0% confidence interval for B	
		B	Std. error	Beta			Lower bound	Upper bound
1	(Constant)	2.459	0.110		22.325	0.000	2.241	2.678
	AvAMU	0.326	0.050	0.533	6.510	0.000	0.227	0.425

[a] Dependent variable: AVPS

influences project success with a B coefficient value of 0.326 and p-value of 0.000 proving that agile method use is statistically significant at the 1% level. The equation of project success in relation to agile methodology use is as follows:

$$AvPS = 0.326AvAMU + 2.5, \tag{2}$$

where AvPS is the project success, AvAMU is an agile method use theorized to likely directly be impacts project success.

This is our main equation for calculating the influence of project success in relation to agile method use. The results discussed above are summarized in Fig. 2 illustrating the influence of individual components of agile methodology use on project success and the influence of agile methodology use as a whole on project success.

5 Discussions of Key Findings

This study aimed to examine the impact that agile method use has on project success in the South African banking sector. This study was motivated by many studies that highlighted that the claims that were made about agile that it improves the project and organization for better have not been proven [24]. A model Fig. 2 was proposed and used to test our theory.

There have been numerous claims made by agile practitioners that agile improves team efficiency, team performance, software quality and organizational benefits [7]. In this study, agile method use had four components, reduced upfront planning aimed at determining the time reduced by the team before beginning the work. Iterative delivery aimed at determining the team taken to deliver the functional work. Environmental feedback aimed at determining the mechanisms utilized by the team to obtain feedback from the customers and stakeholders, and technical feedback aimed at determining the level at which the team used the mechanisms to ensure that the system is functioning properly [33]. These variables were theorized to directly impact project success and proven true.

Project success was measured using project performance, product quality and project impacts. Project performance aimed at determining whether the project results have matched the defined goals of scope schedule and budget. Product quality aimed at determining whether the system is useful, reliable, complete, effective, suitable and accurate as perceived, while project impacts aimed at determining the extent or how positively has the project impacted the organization as perceived [33].

In accordance with [33], our results indicated that the agile method uses positively impact project success. Furthermore, this study has indicated that reduced upfront planning, iterative deliver, environmental feedback and technical feedback as individual agile method use components positively impact project success in the South African banking sector.

6 Conclusion

The results of this research indicated that the agile method use positively impacts project success. In agile method use, the following variables: Reduced upfront planning, iterative delivery, environmental feedback and technical feedback influence project success. In future, it is necessary to understand the effect that the individual agile method use variables (reduced upfront planning, iterative delivery, environmental feedback and technical feedback) have on the individual project success components (project performance, product quality and project impacts). How project success impacts the organizational net benefits indicates a significant need for future research as well.

References

1. Ahimbisibwe A (2015) Critical success factors for outsourced software development projects from a Vendor's perspective: a structural equation modelling analysis of traditional plan-based and agile methodologies. Victoria University of Wellington
2. Altameem E, Mohammad AI (2015) Impact of agile methodology on software development. In: Computer and Information Science. Canadian Center of Science and Education, pp 9–14. doi: https://doi.org/10.5539/cis.v8n2p9
3. Asnawi AL, Gravell AM, Wills GB (2012) Emergence of agile methods: perceptions from software practitioners in Malaysia. Doi: https://doi.org/10.1109/AgileIndia.2012.14
4. Balasundaram N (2009) Factor Analysis: nature, mechanism and uses in social and management science research. J Cost Manage Acc XXXVII(2):15–25
5. Ceschi M et al (2005) Project management in plan-based and agile companies. University of Bolzano-Bozen
6. Chevers DA, Whyte CC (2015) The adoption and benefits of Agile software development methods in Jamaica. In: 2015 Americas conference on information systems, AMCIS 2015, pp 1–9. Available at: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84963616582&partnerID=40&md5=5823bedf6722a72069d185562a3b1e7a
7. Douglass BP, Ekas L (2012) Adopting agile methods for safety-critical systems development. Longman, Britain
8. Fergal M et al (2016) Introducing agile practices into MDevSPICE. Int J Adv Life Sci 8(1):133–142
9. Fergis K (2012) The impact of an agile methodology on software development costs. University of Pennsylvania Technical Papers, pp 1–16. Available at: https://repository.upenn.edu/cis_reports%0A
10. Gaurav Kumar PKB (2012) Impact of agile methodology on software development. Int J Comput Technol Electron Eng (IJCTEE) 2(4):46–50. https://doi.org/10.5539/cis.v8n2p9
11. Gaurav Kumar PKB (2012b) Impact of agile methodology on software development. Comput Inform Sci 8(2). doi: https://doi.org/10.5539/cis.v8n2p9
12. George D, Mallery P (2003) SPSS for windows step by step: a simple guide and reference. 11.0 update, 4th edn
13. Haider A (2017) Impact of agile methodologies on cost estimation techniques in software industry of Pakistan. IndustrEng Manage 6(03). doi: https://doi.org/10.4172/2169-0316.1000218
14. Hneif M, Ow SH (2009) Review of agile methodologies in software development 1. Int J Res Rev Appl Sci 1(1):2076–2734. ISSN:2076-734X, EISSN:2076-7366

15. Kapitsaki GM, Christou M (2015) Learning from the current status of agile adoption. In: Filipe J, Maciaszek L (eds) ENASE 2014, CCIS 551, pp 18–32. Springer International Publishing Switzerland 2015, pp 18–32. doi: https://doi.org/10.1007/978-3-319-27218-4
16. Kautsar E et al (2013) Challenges in adopting agile practices: perceptions of software practitioners in Indonesia. In: The 5th international conference on internet (ICONI) 2013. Malaysia, pp 1–9
17. Lagerberg L et al (2013) The impact of agile principles and practices on large-scale software development projects: a multiple-case study of two projects at Ericsson. Int Symp Empirical Softw Eng Meas 10:348–356. https://doi.org/10.1109/ESEM.2013.53
18. Mahanti A (2006) Challenges in enterprise adoption of agile methods—a survey. J Comput Inform Technol 03:197–206. https://doi.org/10.2498/cit.2006.03.03
19. Masood ZA, Farooq S (2017) The benefits and key challenges of agile project management under recent research opportunities. Int Res J Manage Sci 5(1):20–28. Available at: https://www.researchgate.net/publication/316239082
20. Mcbreen P (2002) Questioning extreme programming. Canada
21. McCrindle C (2008) Choosing and using quantitative research methods and tools. FIL-IDF world dairy summit and exhibition. Available at: https://www.up.ac.za/media/shared/624/choosing-and-using-quantitative-research-methods-and-tools.zp119932.pdf
22. Mchugh M, Fergal M, Coady G (2015) Adopting agile practices when developing medical device software adopting agile practices when developing medical device software. Ireland: creative commons attribution-noncommercial-share alike 3.0 License
23. Nguyen DS (2016) Success factors that influence agile software.pdf. Am Sci Res J EngTechnol Sci (ASRJETS) 17(1):172–222
24. Owen RL, Koskela L (2006) Agile construction project management. In: 6th international postgraduate research conference in the built and human environment, 6(7), pp 22–33. Available at: https://www.irbnet.de/daten/iconda/CIB9021.pdf
25. Penn DM (2016) Agile and conventional methodologies: an empirical investigation of their impact on software quality parameters. IOSR J Econ Finance 3(1):56. https://doi.org/10.3929/ethz-b-000238666
26. Sarstedt M, Mooi E (2014) A concise guide to market research—Chapter 7 regression analysis. Springer Texts in Business and Economics, Berlin. https://doi.org/10.1007/978-3-642-53965-7
27. Serrador P, Pinto JK (2015) Does agile work?—a quantitative analysis of agile project success. Int J Project Manage APM and IPMA 33(5):1040–1051. https://doi.org/10.1016/j.ijproman.2015.01.006
28. Shiner K, Pitt A (2015) Implementing-agile-in-financial-services-intelliware-development.pdf. Toronto, Ontario, Canada
29. Sidky A, Arthur JD (2007) A structured approach to adopting agile practices: the agile adoption framework a structured approach to adopting agile practices: the agile adoption framework. Innov Syst Softw Eng 3(3):1–12
30. Sinha S (2017) Exploratory study on the influence of agile on project management of outsourced software projects in India. Dublin
31. Taber KS (2018) The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ Res Sci Educ 48(6):1273–1296. https://doi.org/10.1007/s11165-016-9602-2
32. Totten J (2017) Critical success factors for agile project management in non-software related product development teams, Dissertation. Western Michigan University. Available at: https://scholarworks.wmich.edu/dissertations/3178
33. Tripp JF (2012) The impacts of agile development methodology use on project success: a contingency view, Michigan
34. Turk D, France R, Rumpe B (2002) Limitations of agile software processes. In: Extreme programming and flexible processes in software engineering, Alghero, Italy, pp 43–46. www.se-rwth.de/publications
35. Vanker C (2015) The adoption of agile software development methodologies by organisations in South Africa. doi: https://doi.org/10.13140/RG.2.1.1831.7683

36. Velumani M (2017) Adoption of agile enterprise architecture in large organisation: a case study
37. VijayasarathyLR, Turk D (2008) Agile software development: a survey of early adopters. J InfTechnol Manage XIX(2):1–8. https://www.aom-iaom.org/jitm_pdfs/jitm_08/article3.pdf
38. Zafar I, Nazir AK, Abbas M (2017) The impact of agile methodology (DSDM) on software project management. In: Circulation in computer science: international conference on engineering, computing & information technology (ICECIT 2017). Islamabad, Pakistan, pp. 1–6. www.ccsarchive.org

Automated Industrial Sound Power Alert System

R. Vishva, P. Harish Annamalai, K. Raja Raman, B. Vijay, J. Rolant Gini, and M. E. Harikumar

Abstract Noise pollution is a threat to health and well-being. Normal environmental noise is around 40–60 dB. But if the noise level increases above 80 dB, it can affect our psychomotor performance. To address this issue, the noise decibel level should be calculated and preventive measures should be taken. This work proposes an Automated sound power alert system at an industrial level that displays the decibel value of noise around heavy machinery and if it exceeds the threshold value of 80 dB, it notifies the authority by delivering an automated text message. It follows message queuing telemetry transport protocol which forms a basis for communication between microcontroller and Adafruit web. Data is transmitted to the web feed using the Message Queuing Telemetry Transport (MQTT) Publish pattern. The web feed is continuously monitored by the IF This Then That (IFTTT) for triggering the alert message. It is also equipped with liquid crystal display, Light-emitting diodes, and Piezo Buzzer that notifies workers around the machinery about the noise level as well as provides continuous monitoring of the same.

Keywords Message queuing telemetry transport (MQTT) protocol · IF this then that (IFTTT) · Liquid crystal display (LCD) · Light-emitting diode (LED)

R. Vishva · P. Harish Annamalai (✉) · K. Raja Raman · B. Vijay · J. Rolant Gini · M. E. Harikumar
Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: harishanna336@gmail.com

R. Vishva
e-mail: vishvar0610@gmail.com

K. Raja Raman
e-mail: 2000rajaraman@gmail.com

B. Vijay
e-mail: vijaybaskar391999@gmail.com

J. Rolant Gini
e-mail: j_rolantgini@cb.amrita.edu

M. E. Harikumar
e-mail: me_harikumar@cb.amrita.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_13

1 Introduction

The word noise is derived from the Latin word “Nausea” implying “unwanted sound” or sound that is loud, unpleasant, or unexpected. It can also be defined as a wrong sound, in the wrong place, and at the wrong time; is more severe and widespread than ever before which is known as noise pollution or sound power. It will continue to increase because of population growth and urbanization [1]. The population in the world continues to increase every day and India is said to become the most populated during 2019 and 2050. This will lead to an increase in noise pollution. Current guidelines used are insufficient to control noise pollution. It has been an underestimated threat that can cause short-term and long-term problems [2]. Continuous exposure to noise leads to hearing impairment and health problems like arterial constriction which leads to high blood pressure, sleep disturbance, etc. Emotional stress triggered by noise was suggested to play a role in the respiratory problems in children [3–5]. In India, the prevalence of hearing loss was estimated to be 63 million (6.3%). Preventive measures of noise pollution include the use of protective equipment like ear muffs and earplugs. Industries must also design machines that produce less noise and use noise abatement equipment such as acoustic louvers, Industrial exhaust silencers [6]. This kind of noiseless industry machinery design requires more cost and it may not be possible to implement in all the industries. The government of India has also implemented Noise Pollution (Regulation and Control) rules which are not regularly implemented and inspected [7]. Many people are not aware of the consequences of noise pollution. A study conducted in Orissa, a state in India found that, though people experienced noise-induced symptoms such as headaches, bad temper, hearing problems, loss of concentration, and sleep disturbance, they were unaware of the ill-effects of noise on health. If every person present in such an environment is aware of the consequences involved and alerted when the environment becomes dangerous then this would lead to the prevention of health effects caused by an increase in noise levels [8]. The current decibel meters are accurate but the user should stay near the heavy machinery for quite an amount of time to take readings. This could lead to hearing impairment and the health problems of the testing person [9].

The proposed Automated Sound Power alert system in this paper uses a condenser microphone to obtain noise input from industrial types of machinery. It is then processed by the microcontroller and the decibel value is displayed on the Liquid Crystal Display (LCD) screen. Moreover, the decibel values that are calculated by the microcontroller are periodically sent to an Adafruit server which acts as an intermediate medium for communication between the microcontroller and IFTTT. The IFTTT reads data from the Adafruit server and triggers the alert message if the value is greater than the threshold value. By employing the proposed system in industries, workers can be made aware of the noise decibel level around their workplace, and also hearing impairment and health problems can be avoided.

The full paper is organized in the following manner. Section 2 describes the design and working of the proposed system. Section 3 represents the results and discussions. Section 4 concludes the research work.

2 Design and Working

In this work, the input signal is obtained from the Condenser microphone, an analog sensor that converts sound pressure into an electrical signal. It is then processed by a preamplifier which converts the weak electrical signal into an output signal strong enough for further processing. The output analog signal is converted to a digital signal with discrete values ranging from 0 to 1024 by the Analog to Digital Converter (ADC) in the microcontroller ESP8266. The resulting digital values are then processed to estimate the noise level by the microcontroller ESP8266 to provide the current situation. LCD and Light-Emitting Diode (LED) are used to display the current decibel level. If the resultant noise level is over the threshold value then the piezo buzzer will alert the workers around. Being a Wi-Fi microchip, the ESP8266 also sends the resultant data to an Adafruit feed. IFTTT, a freeware web-based service is used to read data from the Adafruit server, which helps in triggering the alert message.

The system design is illustrated in Fig. 1, in the form of a block diagram. The Automated sound power alert system has two different ways to alert the workers. It displays the decibel value directly on the LCD screen and the other is the number of LED lights that glow in correspondence to the decibel value. Each LED light that glows denotes a 10 dB noise level in the system. The more the number of light glows, the higher the decibel value. The latter way is very useful to find the decibel value without the user stay near the heavy machinery to find out the value.

The condenser microphone is to be placed between 5 and 10 cm distance from the noise source (Industrial machinery) to be measured for getting good input. The

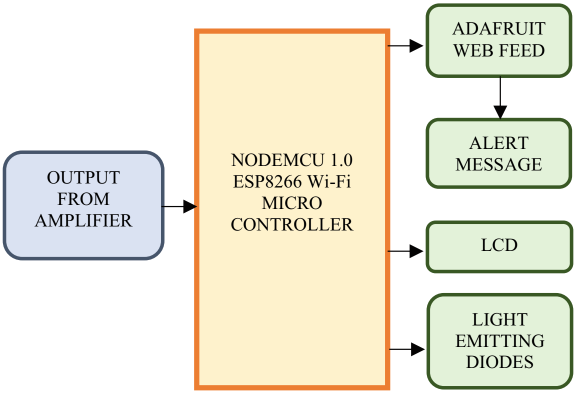


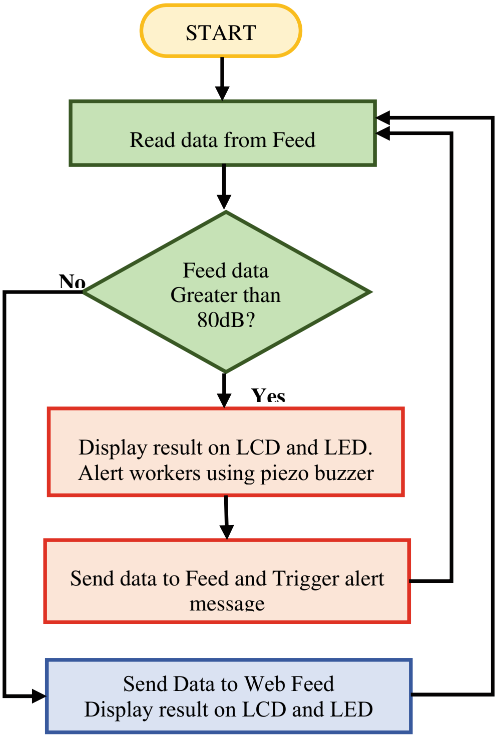
Fig. 1 Block diagram of the system

preamplifier amplifies the signal, which is then processed by the microcontroller and the result is displayed. The microcontrollers can be powered using micro USB 2.0 cables. Using the MQTT library, the data has been transmitted periodically from the microcontroller to the web feed. Adafruit is a free web-based service where only 30 data points can be sent to the server for a minute. Hence, the Automated sound power alert system sends the value to the feed every two seconds. The IFTTT reads data from the updated feed from the Adafruit web server and triggers the alert message when the value is over the threshold value. The basic control flow of the Automated sound power alert system is illustrated in Fig. 2.

The 16X2 LCD used in this experiment requires 5 V DC voltage to display output. Since the General Purpose Input–Output pins (GPIO) in ESP8266 microcontroller provide only 3.3 V. The V_{in} pin of the ESP8266 microcontroller gives 5 V DC. So, an IC7408, commonly known as AND gate is used to get 5 V from V_{in} pin to power up the LCD. From Fig. 3, the circuit connection for LCD through IC7408 is understood.

Inputs obtained from the condenser microphone may not be strong enough for processing. It needs to be amplified to convert it into a digital signal. The LM358, a preamplifier shown in Fig. 4b, is used to amplify the weak input signal. By using Preamplifier, it is possible to obtain a higher gain value and improved sound quality. The LEDs are connected to GPIO pins of the ESP8266 Wi-Fi Microcontroller. From

Fig. 2 Control flow diagram of the proposed system


START
Read data from Feed
Feed data Greater than 80dB?
No
Yes
Display result on LCD and LED. Alert workers using piezo buzzer
Send data to Feed and Trigger alert message
Send Data to Web Feed Display result on LCD and LED

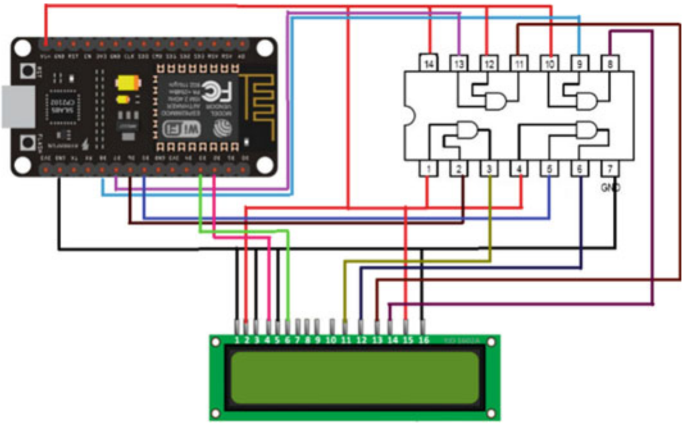


Fig. 3 LCD circuit connection of the proposed system

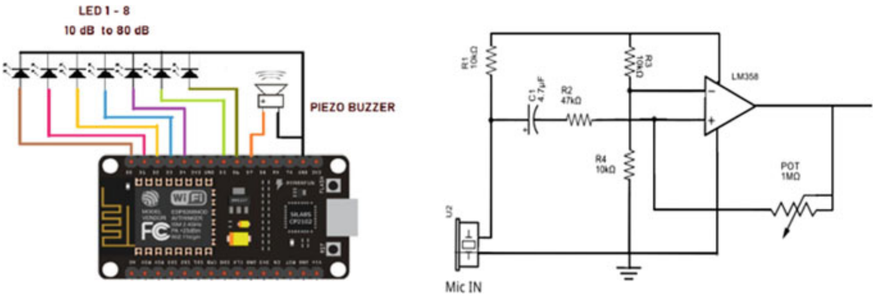


Fig. 4 **a** LED circuit connection of the proposed system. **b** Preamplifier circuit connection

Fig. 4a, the circuit connection for LEDs and From Fig. 4b, preamplifier circuit connection is understood. The components used and their purpose in this Automated sound power alert system are given in Table 1.

Table 1 Components used in the automated sound power alert system

Components	Purpose
Condenser microphone	The most common instrumentation microphone—used to get input signals
LM358 (preamplifier)	To amplify the weak signal to a strong output signal for processing [10]
ESP8266 Wi-Fi microchip	The processing unit for the sound power alert system [11]
IC7408 and LCD	Display the decibel value for the user
LEDs and Piezo Buzzer	To indicate the decibel range for the user

2.1 *Software Used*

Arduino IDE is an open-source Arduino software [12] is used for control code and uploaded to the ESP8266 Microcontrollers. Adafruit IO, which supports MQTT protocol [13], acts as a feed for the data transmitted by the ESP8266 microcontroller. IFTTT, a free web-based service [14] is used to monitor the changes at Adafruit web feed and trigger an alert message if the feed value is over the threshold value.

3 Results and Discussion

The Automated Sound power alert system is tested for various noise inputs. For experimental trails, inputs are given through a Bluetooth speaker. Input for the speaker consists of various songs that are played at different volumes and also inputs are taken from a very noisy traffic environment. The outputs are displayed on LCD and LEDs. The interpretation of decibel values from the LEDs is understood from Table 2.

The working model of the sound power alert system can be understood from Fig. 5. As soon as the condenser microphone receives noise from various sources, it converts it into an electrical signal and the preamplifier amplifies it. The processed result is displayed in the LCD and LEDs. The number of LEDs that glow, in this case, 5 corresponds to the range 50–59 dB which is also shown by the LCD.

The data which is displayed in the LCD, in this case, 57 dB is sent to the Adafruit web feed. The same has been shown in Fig. 6a. IFTTT provides many web-based services. These services are known as applets. For the proposed automated sound power alert system, the Adafruit applet with the android SMS applet is combined to trigger alert messages. Since ESP8266 doesn't have an inbuilt messaging service this is setup is required. The IFTTT setup interface has also been shown in Fig. 6b. For Adafruit feed values that are over the threshold value, the IFTTT triggers an alert message.

The input to the Bluetooth speaker is given at maximum volume. From Fig. 7a, it is understood that the calculated decibel value is over the threshold value of 80 dB

Table 2 Decibel range correspondence

Decibel range	Number of LED glows
10–19 dB	1
20–29 dB	2
30–39 dB	3
40–49 dB	4
50–59 dB	5
60–69 dB	6
70–79 dB	7
Greater than 80 dB	7 + 1 (Piezo Buzzer)

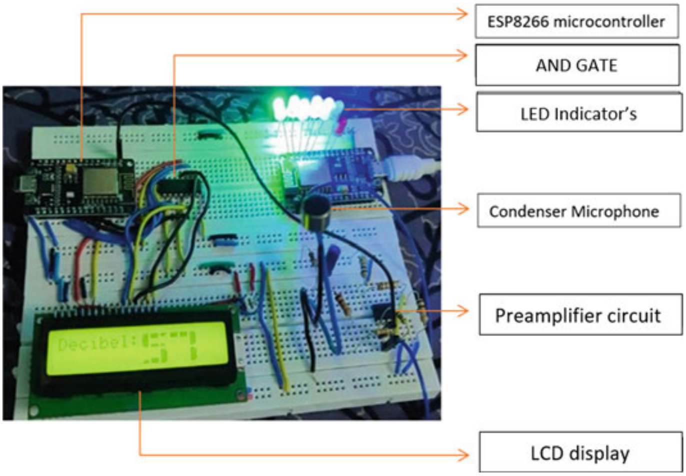
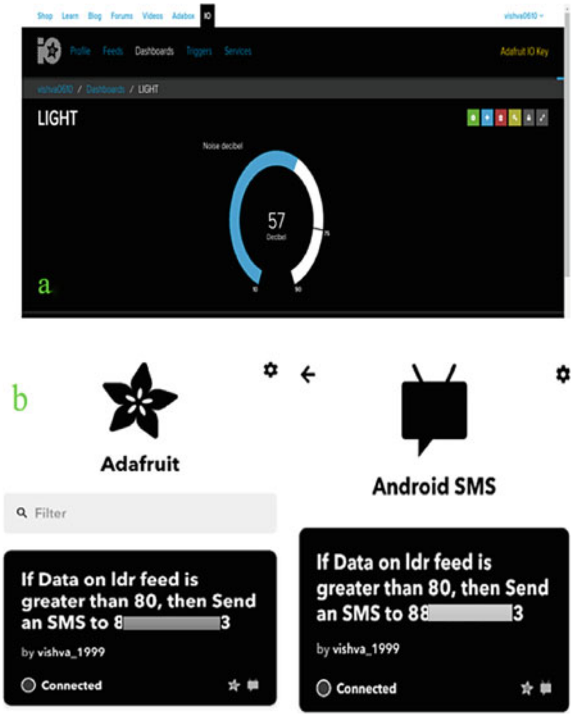


Fig. 5 Working model of the automated sound power alert system


Shop Learn Blog Forums Videos Adabox IO
vishva0610
Profile Feeds Dashboards Triggers Services
Adafruit IO Key
vishva0610 / Dashboards / LIGHT
LIGHT
Noise decibel
57
Decibel
a
b
Adafruit
Filter
Android SMS

If Data on ldr feed is greater than 80, then Send an SMS to 8[redacted]3

by vishva_1999

○ Connected

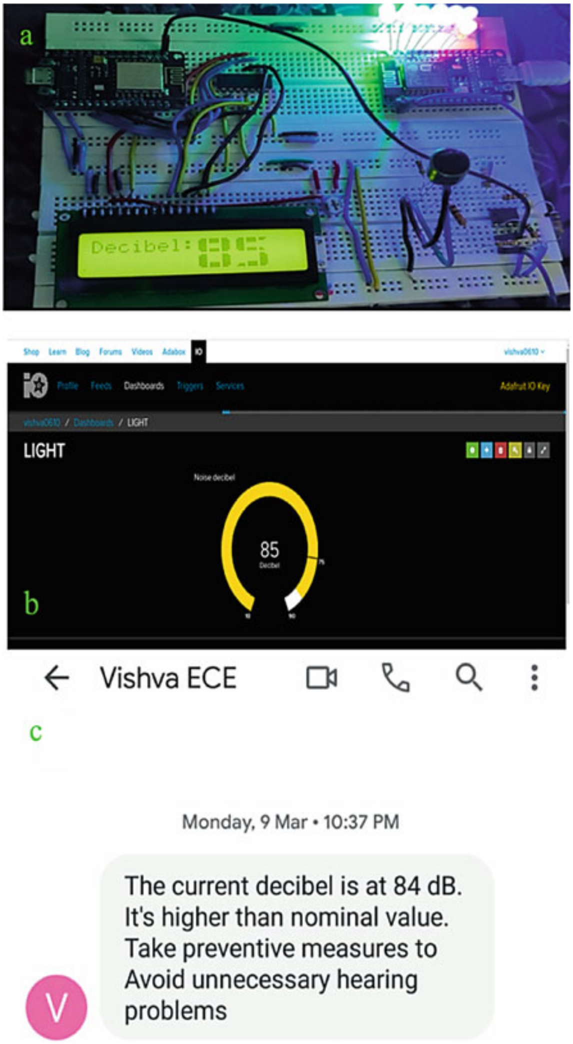
If Data on ldr feed is greater than 80, then Send an SMS to 88[redacted]3

by vishva_1999

○ Connected

Fig. 6 **a** Adafruit Web Feed displaying transmitted data from ESP8266 Wi-Fi Microcontroller. **b** IFTTT setup interface for Adafruit and Android SMS

Fig. 7 **a** Sound power alert system result for the noisy case. **b** Adafruit Web Feed displaying transmitted data. **c** The triggered alert message by the IFTTT



Monday, 9 Mar • 10:37 PM

The current decibel is at 84 dB. It's higher than nominal value. Take preventive measures to Avoid unnecessary hearing problems

and the same has been displayed in the LCD. The MQTT publish protocol transmits the resulting value which is displayed on the LCD to the Adafruit web feed. The transmitted value is shown on the Adafruit feed dashboard as shown in Fig. 7b. The IFTTT, which connects the Adafruit and the android SMS applets monitor the Adafruit feed, finds the value from feed to be greater than the threshold value of 80 dB. so, it triggers an alert message as programmed. The alerted text message is as displayed in Fig. 7c.

The inputs given to the condenser microphone should be monotonous. Calibration of the condenser microphone is a very important process. Since Adafruit feed can support only 30 data points per minute, data is transmitted to the Feed with a significant delay of 2 s. Care must be taken so that the two ESP8266 controllers aren't out of sync. This is the drawback of the proposed Automated sound power alert system. The use of a Wi-Fi microcontroller with a greater number of General Purpose Input–Output (GPIO) pins to support both LCD and LEDs at the same time could solve the above case. This can be upgraded and made to work collectively with other finished products. For example, the Police of Mumbai city, India has equipped traffic signals with noise decibel meters, which increases the waiting time for the signal to turn green if the decibel value crosses 85 dB [15, 16]. This reduces people from honking unnecessarily while waiting for the traffic signal. This can also be extended for domestic use with minor changes.

4 Conclusion

The proposed Automated sound power alert system is lightweight, low cost, and user friendly for workers to use. It is easy to use and it does not require a noise assessor to be present all the time. Health risks like stress and hearing loss can be minimized to an extent with the use of these systems. It helps industries in managing Occupational Health and Safety claims regarding the noise of machineries. The proposed Automated sound power alert system would create a healthy work environment that improves productivity at an affordable cost. This would increase the Gross Domestic Product and economy of a country.

References

1. Singh N, Davar SC (2004) Noise pollution-sources, effects, and control. J Hum Ecol 16:181–187
2. Smys S, Raj JS (2019) Virtual reality simulation as therapy for posttraumatic stress disorder (PTSD). J Electronic 1(1):24–34
3. Hahad O, Kroller-Schon S, Daiber A, Munzel T (2019) The cardiovascular effects of noise. Deutsches Arzteblatt International 116:245–250
4. Clark WW, Bohne B (1984) The effects of noise on hearing and the ear. Medical Times 112:17–22
5. Petric D (2020) Detrimental health effects of noise pollution
6. Mohanapriya SP, Sumesh EP, Karthika R (2014) Environmental sound recognition using Gaussian mixture model and neural network classifier. In: International conference on green computing communication and electrical engineering (ICGCCEE), Coimbatore, pp 1–5
7. High Court warns of action for non-implementation of noise pollution rules. https://economictimes.indiatimes.com/news/politics-and-nation/high-court-warns-of-action-for-non-implementation-of-noise-pollution-rules/articleshow/47800925.cms
8. Balazikova M, Salaj L, Wysoczanska B (2019) Analysis of human factor reliability in workplace with noise load. In: 2019 international council on technologies of environmental protection (ICTEP), Stary Smokovec, Slovakia, pp 25–29
9. Davis AH (1938) An objective noise-meter for the measurement of moderate and loud, steady and impulsive noises. J Inst Electr Eng 83(500):249–260

10. Neri B, Pellegrini B, Saletti R (1991) Ultra-low-noise preamplifier for low-frequency noise measurements in electron devices. In: IEEE transactions on instrumentation and measurement 40:2–6
11. Raj JS, Vijitha Ananthi J (2019) Automation using IoT in greenhouse environment. J Inform Technol 1(1):38–47
12. Ramon MC (2014) Arduino IDE and wiring language. In: Intel® Galileo and Intel® Galileo Gen 2. Apress, Berkeley, CA
13. Jayan AP, Balasubramani A, Kaikottil A, Harini N (2019) An enhanced scheme for authentication using OTP and QR code for MQTT protocol. Int J Recent Technol Eng 7:70–75
14. Ovadia S (2014) Automate the internet with "If This Then That" (IFTTT). Behav Soc Sci Libr 4:208–211
15. Mumbai Police Play a Trick on Honking Drivers. https://www.nytimes.com/2020/02/04/world/asia/mumbai-horn-honking.html
16. The more you honk, the longer you wait! Signal to stay red if decibel level high. https://timesofindia.indiatimes.com/city/mumbai/more-you-honk-the-longer-you-wait-signal-to-stay-red-if-decibel-level-high/articleshow/73819813.cms

Predicting Early Readmission of Diabetic Patients: Toward Interpretable Models

Mir Moynuddin Ahmed Shibly, Tahmina Akter Tisha, and Md. Mahfuzul Islam Mazumder

Abstract Hospital readmission among diabetic patients is a common phenomenon throughout the world. Predicting such patients with a high risk of readmission at the time of discharge even before can help us to provide better health care to them. It can minimize the cost associated with readmission too. This study aims at creating a decision support system that can find diabetic patients who are prone to early readmission. To do that, several data mining techniques have been used. Two regular classifiers using the decision tree and random forest have been developed. After that, two rule-based classifiers using Repeated Incremental Pruning to Produce Error Reduction (RIPPER) and PART algorithms have been developed to provide better interpretability and understandability of the support system. Between two regular classifiers, the random forest has shown a better performance with 89.5% accuracy and 89.5% recall. And, between rule-based classifiers, PART has demonstrated promising performance with 86.6% accuracy and 84.6% recall. Using these classifiers, smart and improved health care can be ensured.

Keywords Rule-based classifiers · RIPPER · PART · Decision support system · Dataset imbalance · Synthetic Minority Oversampling Technique (SMOTE)

1 Introduction

Hospital readmission is an important issue in the health care system. It can be a matter of inconvenience for patients, doctors, and other stakeholders. It can also unbalance the overall cost management of the health care system. Around 17 billion dollars is spent yearly on hospital readmission in the USA [1]. Follow-up treatment

---

M. M. A. Shibly (✉) · T. A. Tisha · Md. M. I. Mazumder
Department of Computer Science and Engineering, East West University, Dhaka, Bangladesh
e-mail: shiblygnr@gmail.com

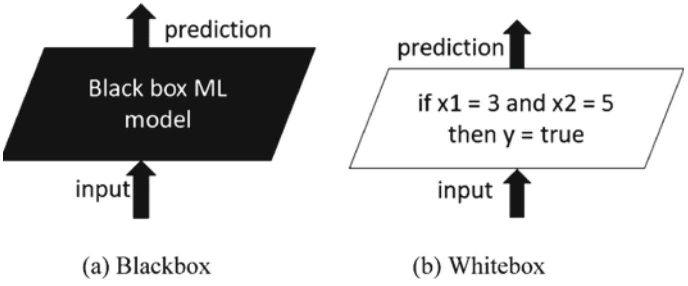
T. A. Tisha
e-mail: tahminatish001@gmail.com

Md. M. I. Mazumder
e-mail: 2016-3-60-048@std.ewubd.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_14

of the critical patients and taking appropriate measures can reduce the rate of readmission. The readmission can happen because of various reasons like heart failure, pneumonia, acute myocardial infarction, etc., [2]. But among the patients who have stayed in a hospital for various reasons, diabetic patients are more vulnerable to hospital readmission [3]. The reason for diabetic patients to stay in the hospital does not necessarily have to be a diabetes-related illness. Readmission often occurs for those patients who have diabetes as comorbidity but admitted for other reasons. A study shows that the readmission of diabetic patients can be managed by better follow-up treatment [4]. To ensure better treatment after discharge, the first step is to identify those who are at risk of getting readmitted early. A statistical decision support system based on data mining techniques that can detect those risky patients early can be helpful to mitigate the risks of mortality. Early predicting of them using this system can also be financially helpful. In the current world where ensuring proper health care for people is the most challenging task; such support systems can make life easier for all stakeholders. If a patient who is likely to be readmitted early after discharge, then he/she can be monitored with high importance to prevent readmission. Customized health care plan can be designed for the targeted patients. Such precautions can result in to provide better health care services in terms of reducing risk of getting readmitted and improving cost management associated with it.

For being an important and challenging issue in the health care system, many methods have been used to accomplish the task of early prediction of readmission of diabetic patients. Different researchers have applied different techniques to provide classification-based predictive models. Some have used machine learning algorithms [5] like support vector machine [6], random forest [7], neural networks [8], etc. The typical machine learning models can act as a black box, i.e., how the model decides something based on input is unknown to the users. Figure 1a illustrates an example of such a machine learning model. Usually, the neural network-based models are treated as the black box system. If the users cannot perceive how a model makes a decision, then there might be a chance of growing untrust among them [9]. In health care-related problems, the predictions must have to be interpretable, understandable, and transparent. Few methods interpret the results of a machine learning model like Local Interpretable Model-agnostic Explanations (LIME) [10].



(a) Blackbox (b) Whitebox

Fig. 1 Machine learning models

In this domain of the health care system, rule-based classifiers can be utilized to design an improved predictive support system. As their results are easier to understand than the other "black box" systems, rule-based classifiers can also provide better interpretability of the system to the stakeholders. Figure 1b illustrates a rule-based predictive system. There is a gap of knowledge of how the rule-based classifiers can perform in early readmission prediction of diabetic patients. There have not been many works to predict early readmission using rule-based interpretable classifiers.

This study designs interpretable rule-based classifiers to predict the early readmission possibility of diabetic patients. The objectives of this study are:

1. To predict early readmission of diabetic patients using traditional classifiers.
2. To develop interpretable rule-based predictive models using Repeated Incremental Pruning to Produce Error Reduction (RIPPER) and PART algorithms.
3. To compare the results of traditional and rule-based classifiers.

This article is arranged in the following manner, Sect. 2 presents the related works to this study, and Sect. 3 describes the methodologies and datasets. Section 4 presents the results of this study with appropriate comparative analysis, and Sect. 5 discusses the results. And, finally, Section 6 draws some concluding remarks.

2 Related Works

Cui et al. [6] have conducted experiments on data of diabetic patient hospital readmission. The aim has been to reduce the readmission rate so that better health care can be provided by mitigating associated costs incur for hospital readmission. With the samples from the hospital readmission dataset [11] based on a criterion, a dataset has been constructed by the authors. They have used synthetic minority oversampling technique to reduce the imbalance between the class that has early hospital readmission risk and the class that has less chance of getting readmitted. They have used a hybrid feature selection method combining the filter method and wrapper method. For the binary classification task, they have proposed a Support Vector Machine-based (SVM) method. They have used a genetic algorithm to optimize the parameters of SVM. They have also employed k-fold cross-validation in the training stage. The study has achieved 81.02% accuracy in the testing phase.

In another study [7], an intelligent decision support system has been developed using the random forest algorithm and Bayesian network. They have experimented with a series of classification algorithms, and the best performing classifiers have been selected. Before training, they have created a dataset by taking samples from the hospital readmission dataset [11] based on diabetes medication. They have achieved up to 82.97% accuracy with their works. Additionally, they also have identified the optimal medications for 28 comorbidity combinations to reduce the chance of readmission after discharge by increasing monitoring for risky patients.

Another study has been conducted by Bhuvan et al. [12]. They have used naive Bayes, Bayes network, random forest, and neural networks for the same classification problem. They have obtained a maximum of 65.4% accuracy. Additionally, they have calculated the costs that can be saved by employing their models. Salian et al. [13] have used traditional classifiers decision tree, logistic regression, SVM, k-nearest neighbors to accomplish a similar task. They have achieved a minimum misclassification rate of 28% using a decision tree. Additionally, they also have generated some rules from the decision tree. Similar studies have been conducted by Alajmani et al. [14], Alloghani et al. [5], Shameer et al. [15], etc. Different studies have utilized different techniques but to achieve a common goal. All of these studies are aimed to find predictive systems that can identify the high-risk patients that are more likely to get readmitted to the hospital 30 days after discharge. Early identification of such patients can help the stakeholders by following up with better treatment and by minimizing cost. After analyzing these related works, it has been seen that there have not been enough significant works using rule-based classifiers to predict early readmission.

3 Methodology

The objective of this study is to predict the risk of a diabetic patient being readmitted to the hospital 30 days after discharge. This work designs a decision support system to predict whether a diabetic patient is needed to be under constant monitoring or not after discharge. In this section of the article, the methods and materials that have been used to create such a support system are described. The complete flow of the works has been demonstrated in Fig. 2.

3.1 Dataset

In this study, “Diabetes 130-US hospitals for years 1999–2008 dataset” [11] containing more than 100,000 instances from the UCI machine learning repository has been used. This dataset has been constructed from the health facts database (Cerner Corporation, Kansas City, MO) based on five criteria. There are 50 features like age, race, weight, the medical speciality of a patient, different diagnoses and medications, etc., in this dataset. Each instance of the dataset is labeled with a class having one of the three outcomes—not readmitted, readmitted in less than 30 days, and readmitted after 30 days. Eleven of the features of the dataset are numeric, and the rest are nominal.

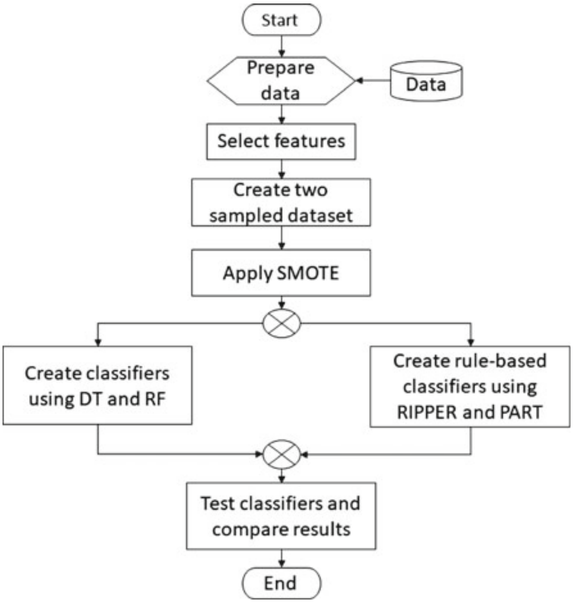

Start
Prepare data
Data
Select features
Create two sampled dataset
Apply SMOTE
Create classifiers using DT and RF
Create rule-based classifiers using RIPPER and PART
Test classifiers and compare results
End

Fig. 2 Flow diagram of the study

3.2 Data Preparation and Feature Selection

There are few nominal features in the dataset that contain null values. The features that have less than or equal to 2% missing values are replaced with the mode of the corresponding feature. On the other hand, the feature "medical specialty" has 53% missing values. They are replaced with a new feature value named "unknown." Another feature named "weight" which has 97% missing entries is removed from the dataset. Three other attributes—"encounter ID," "patient no", and "payer code" are also removed. Three categorical variables "diag_1," 'diag_2," and "diag_3" contain international statistical classification of diseases and related health problems, i.e., ICD9 codes of primary, secondary, and tertiary diagnosis of a patient. These features are mapped to nine major diagnoses as suggested in the original paper of this dataset [11]. In this study, the focus group is the patients with readmission status "<30." The ">30" status has been merged with the "NO" group to make it a binary classification problem. In this way, a complete categorical dataset with 45 features with the target feature having binary classes is prepared. The dataset has been needed to be binarized further in some experiments for the compatibility issue of used tools. Additionally, for feature selection, the Gini index—a filter model has been used. Gini index calculates the discriminative power of each feature [16]. Using this, 24 features with lower scores are selected for classification.

3.3 Data Sampling

The initial experiments of this study have been conducted on the complete dataset. But our scope is only diabetic patients. Two sampled datasets have been prepared based on two criteria. With the data points that have a primary diagnosis as diabetes, i.e., diag_1 = Diabetes and diabetes_med = Yes, two sampled datasets are created as suggested in these papers [6, 7]. In the original dataset and two sampled datasets, there is an imbalance between two target classes. This imbalance in datasets may result in poor performance of the classifiers. And, the classifiers can perform poorly to detect the patients who have a higher chance to get readmitted despite performing well overall. To tackle this imbalance, the Synthetic Minority Oversampling Technique (SMOTE) [17] has been used. In this method, each instance of the minority classes' k-nearest neighbors from the same class is calculated. After that, a portion of the neighbors is selected randomly based on the number of oversampling instances needed. Then, for each instance neighbor pair, a synthetic data point is generated and added to the training set.

3.4 Classifiers

After preparing data and feature selection, two classifiers have been developed using decision tree and random forest algorithms. The decision tree algorithm continuously divides the work area by plotting lines into sub-areas until a specific class emerges. In the decision tree algorithm, the process of classification is created with a set of hierarchical decisions on the attributes structured with a tree [16]. The decision at a node is a condition on one or more attributes in the training set. The condition divides the training data into two or more parts. The aim is to separate the training data into a smaller portion of the target classes in the best possible way. In this manner, the tree grows until it meets a stopping criterion. A stopping criterion can be where all the training examples in a leaf node belong to a single class. This algorithm finds the best split using some quantifications. In this study, the decision tree with a Gini index has been used for classification tasks.

Another classification algorithm that has been used is the random forest algorithm. It is an ensemble learning method which is widely used in regression and classification problems. A random forest is nothing but a combination of many decision trees. In this method, k decision trees are created using training data to form a forest, and for the test data, the majority voting technique is followed. The majority class predicted by individual decision trees is the predicted class. The term forest came from the idea that to create each decision tree, a random set of attributes is selected. The trees are not developed with all features. Another way to create a random forest to use a bootstrap aggregating technique [18]. A portion of the training data points is selected with replacement, and for each decision tree, a newly sampled training dataset is used. Random forest resolves the problem of overfitting in decision tree classifier [19]. In this study, random forest with random attribute selection has been used.

3.5 Rule-Based Classifiers

As the objective of this study is stated, in this stage, three rule-based classifiers have been developed. The rule-based classifiers have better interpretability than the typical classifiers. How a regular trained model predicts a class in a real-life environment can be less understandable by the common stakeholders who use the system. But if there is some rule upfront supporting a classifier, then the working mechanism behind a decision can be perceived by the patients, doctors, and the other stakeholders. Rules can be generated from a decision tree. But the number of rules generated by a decision tree is enormous. That is why algorithms specially designed for rule learning are used. Three algorithms have been used to generate rules, namely Apriori algorithm, Repeated Incremental Pruning to Produce Error Reduction (RIPPER), and PART algorithm.

3.5.1 Apriori

Apriori is an association pattern mining algorithm that generates frequent itemsets present in a dataset. This algorithm generates a smaller candidate itemsets first. Then, the support of each itemset is calculated. By joining the itemsets that satisfy minimum support, the next smaller candidate itemsets are generated rather than generating all itemsets. From these candidate itemsets, the downward closure property violating itemsets are removed. After that, the supports of the remaining candidates are calculated, and the satisfying itemsets are added to the set of frequent itemsets. This process continues until all the frequent itemsets based on minimum support are found. After mining frequent itemsets, a suitable rule generation framework can be employed which works based on conditional probability. Only the rules that have greater or equal to minimum confidence are added to the ruleset. In this study, a modified rule generation framework has been used. Rules are being looked at that can predict the classes. First, the frequent itemsets that have either of the two classes are filtered. After that, only the rules of types $\{x_1x_2\ldots x_n\} \rightarrow \{readmitted = '{<}30'\}$ and $\{x_1x_2\ldots x_n\} \rightarrow \{readmitted = 'NO'\}$ are considered for association rules, where $x_1x_2\ldots x_n$ are individual items in a frequent itemset. If these rules satisfy the minimum confidence, then they are the desired rules mined.

3.5.2 Repeated Incremental Pruning to Produce Error Reduction

RIPPER is a sequential covering algorithm. It adds three improvements over the Incremental Reduced Error Pruning (IREP) algorithm. IREP starts pruning the created rule right after it was developed. In basic IREP, the training dataset is divided into two sets naming “growing set” and “pruning set.” The growing set has two-third of the training data, and the pruning set has the other one-third. For the growing set, using a basic sequential covering algorithm, the best rule is generated. After that, the worth of the rule is calculated on the pruning set based on a measurement metric. Then, a clause from the end of the rule is omitted, and the worth of the reduced rule is computed. If the worth does not decline, then the pruning process continues. After a rule is pruned, the training instances covered by the rule are removed, and the next rule generation starts. In this way, the classification rules are generated [20]. RIPPER offers few modifications in the IREP procedure. The first improvement is a better metric to calculate the worth of a rule [21]. If the total number of positive and negative class in the pruning set is P and N, and p and n are the number of positively and negatively identified instances, respectively, then the worth of a rule is defined by $(p + (N - n))/(P + N)$. The modification made by RIPPER is to use $(p - n)/(p + n)$ as the metric.

Another limitation of IREP is that it stops adding new rules when the last rule has more than 50% error on the pruning set. The developer of RIPPER has said that this stopping criterion is too early and has suggested a new stopping criterion based on description length. Description length of a rule is the number of bits needed to encode a rule, and the number of examples covered by the class-specific rule in the training dataset, which belongs to a different class, i.e., error made by the rule. After generating a rule, its description length is computed, and if that is b bits larger than the smallest description length obtained so far, then it is not added to the ruleset. The last improvement is rule optimization using minimum description length. After creating all the rules, each of them is analyzed, and two variants of it are created—one is the extended version, and the other is generated from scratch. Among the three, a rule is selected according to the minimum description length. In this study, the RIPPER algorithm has been used to create a rule-based classifier.

3.5.3 PART

The last algorithm to create a rule-based classifier is PART. It is developed by Frank et al. [22]. As mentioned earlier, RIPPER globally optimizes the rulesets after rule generation using complex methods. In contrast, PART does not need this global optimization step to generate accurate rules. It uses a typical separate-and-conquer method to build a rule. It follows a sequential covering algorithm. But it differs in how a rule is created. To create each rule, it creates a partial decision tree on current instances. After building the partial tree, one of the leaves that have the highest coverage of training instances is turned into a rule. In this method, the global optimization step is not needed. Even without this optimization step, the algorithm can perform well. This method has been used to build the last classifier to predict early hospital readmission among diabetic patients.

4 Results

This study aims to develop a system that predicts early hospital readmission of diabetic patients. After data preparation, decision tree and random forest classifiers have been developed on the original dataset including 44 categorical attributes with a target feature having binary classes. Then, 24 features have been filtered using a Gini index. Then, two sampled datasets are created based on primary diagnosis and diabetes medication. The same experiments have been carried out on these two datasets too. In the final phase of the study, three rule-based predictive models have been developed using Apriori, RIPPER, and PART algorithms. To implement the decision tree and random forest classifiers, the scikit-learn library from Python has been used. And, for the rule-based models, Weka—a data mining tool developed at the University of Waikato, New Zealand has been used. For all the classifiers, the datasets have been split into training and testing set having 80 and 20% data, respectively. In this section of the article, the outcome of the study based on some evaluation metrics, and the prominent rules generated by predictive models have been presented. All the performances that are reported in this section are test performances.

4.1 Evaluation Metrics

Four evaluation metrics have been used throughout the whole working process of this study—precision, recall, f1-score, and accuracy. Precision is the portion of the correct predictions made by the classifier with respect to total predicted classes. And, recall is the portion of the correct predictions made by the classifier regarding the total existing accurate classes. F1-score is the harmonic mean of precision and recall, i.e., both precision and recall are given equal importance while evaluating the performance of a classifier. And, accuracy is nothing but the portion of the classifier being right.

$$\text{precision} = \frac{TP}{TP + FP} \tag{1}$$

$$\text{recall} = \frac{TP}{P} \tag{2}$$

$$f1\_\text{score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{3}$$

$$\text{accuracy} = \frac{TP + TN}{P + N} \tag{4}$$

Here,P, N TP, TN, FP, and FN are the number of actual positive instances, negative instances, true positive, true negative, false positive, and false negative, respectively.

4.2 Results Analysis

The initial experimentations on the full dataset show that the overall performance of decision tree (DT) and random forest (RF) is good. The accuracies of classifiers are 88.46 and 89.24%, respectively. But we are more interested to detect the patients that have a higher chance of readmission than the patients that have no chance of readmission. Both classifiers have performed poorly to find out the patients that have been readmitted within 30 days after discharge. But after adding synthetic data using SMOTE, the class-wise performance has been improved. Table 1 shows the comparison of the classifiers' class-wise performance. Another aspect is to consider the impact of feature selection. It has been observed that the classifiers with 45 features and with 24 features have similar performance. This justifies the employing feature selection strategy before training. A comparison with and without feature selection has been demonstrated in Table 2.

The scope of this study is focused on diabetic patients. To provide a better decision support system for this domain-specific patient, the original dataset has been sampled into two sub-samples. In the original paper of the dataset [11], it has been said that the probability of hospital readmission depends on the primary diagnosis of the patients

Table 1 Class-wise performance of the classifiers on the full dataset with and without SMOTE

Sampling technique	Algorithm	Classes	Precision	Recall	Support
Without sampling	DT	<30	0.30	**0.05**	2189
		NO	0.90	0.98	18,165
	RF	<30	0.48	**0.01**	2189
		NO	0.89	1.00	18,165
SMOTE	DT	<30	0.86	**0.90**	17,930
		NO	0.90	0.85	18,234
	RF	<30	0.98	**0.91**	17,930
		NO	0.92	0.99	18,234

Table 2 Impact of feature selection with Gini index

Algorithm	Without feature selection			With feature selection		
	Precision	Recall	Accuracy	Precision	Recall	Accuracy
DT	0.884	0.883	**0.883**	0.883	0.882	0.882
RF	0.952	0.950	0.950	0.952	0.950	**0.951**

Table 3 Performance of decision tree and random forest on two sampled datasets

Dataset	Algorithms	Precision	Recall	F1-score	Accuracy
Sampled based on the primary diagnosis	DT	0.792	0.793	0.792	0.792
	RF	**0.896**	**0.895**	**0.895**	**0.895**
Sampled based on diabetes medication	DT	0.795	0.797	0.796	0.797
	RF	**0.906**	**0.906**	**0.906**	**0.906**

(feature name: diag_1). The ICD9 codes ranging from 250 to 251 denote that a patient is diagnosed with diabetes. The data points that have diag_1 value in this range are considered for the experiment in this phase of the study. 8772 data points fall under this criterion. In the next phase, the data points that have diabetes medication equal to "Yes" has been considered as the samples. 78,363 data points fulfill the diabetes medication criterion. 24 features selected in the previous phase of the study have been used in these experiments. Since synthetic minority oversampling technique has been proven to work better to tackle both imbalance and overfitting problems, it has been used in the experiments too. In Table 3, the performances of the classifiers created on these two sampled datasets have been shown. For both datasets, random forest classifiers have outperformed the decision tree classifiers. The decision tree classifiers have an accuracy of 79.2 and 79.7% in two datasets, respectively, while the random forest classifiers have an accuracy of 89.5 and 90.6%, respectively. And, the class-wise recalls of the classifiers have also been observed. Since SMOTE has been applied on the dataset before training, unlike the initial experiments, the recalls of individual classes have been better.

The results show that classifiers to predict hospital readmission of diabetic patients can be developed using decision tree and random forest. But this type of classifier lacks interpretability. To provide a better decision support system for the doctors, patients, and other stakeholders, rule-based classifiers have been developed. The predictive rules have been generated using Apriori, RIPPER, and PART algorithms. While creating rules from frequent itemsets that are mined using Apriori, only the itemsets having either of two target classes are selected. For mining frequent itemsets, minimum support of 0.2 is used. And for rules have been generated with a minimum confidence of 0.7. In Table 5, five rules generated using the Apriori algorithm are demonstrated. In this study, two rule-based classifiers have also been developed to provide the interpretability of the models. The algorithms used for rule-based prediction are RIPPER and PART. These two algorithms have produced 90 and 635 rules, respectively. Table 4 shows a few of the top rules generated by two algorithms.

The rules generated by three predictive models are easy to understand. The patients and doctors will understand the reason behind a decision that has been made by just looking at them. The rules are also self-explanatory. If the first rule generated by the Apriori algorithm (Table 5) is looked at, then it will eventually be apparent that a patient who has not been under any procedure (*num_procedures* = *(−2, 0]*), who has not any emergency (*number_emergency* = *(−10, 0]*), and has not any inpatient visit (*number_inpatient* = *(−5, 0]*) during the stay in hospital is not likely to be readmitted.

Table 4 Rules generated by RIPPER and PART

No	RIPPER	PART
1	(medical_specialty = unknown) and (number_inpatient = (0, 5]) and (num_lab_procedures = (40, 60]) and (race = Caucasian) and (A1Cresult = None) and (diag_3 = Circulatory) and (diag_2 = Circulatory) → readmitted = <30	number_inpatient = (10, 15] AND medical_specialty = unknown AND A1Cresult = None → readmitted = <30
2	(medical_specialty = unknown) and (num_lab_procedures = (40, 60]) and (race = Caucasian) and (A1Cresult = None) and (diag_3 = Other) and (admission_type_id = Emergency) → readmitted = <30	glipizide = Up → readmitted = NO
3	(number_inpatient = (0, 5]) and (medical_specialty = unknown) and (age = [40–50)) and (diag_2 = Neoplasms) → readmitted = <30	number_inpatient = (5, 10] AND admission_source_id = AS1 AND race = Caucasian AND medical_specialty = unknown → readmitted = <30
4	(number_inpatient = (0, 5]) and (medical_specialty = InternalMedicine) and (number_emergency = (0, 10]) and (age = [40–50)) → readmitted = <30	number_inpatient = (5, 10] AND admission_source_id = AS7 AND discharge_disposition_id = D1 AND race = Caucasian AND diag_3 = Other AND num_medications = (10, 20] → readmitted = <30
5	(metformin = Steady) and (num_lab_procedures = (20, 40]) and (admission_source_id = AS7) and (number_diagnoses = (5, 10]) and (number_emergency = (-10, 0]) → readmitted = NO	glipizide = Steady AND number_emergency = (-10, 0] AND admission_source_id = AS7 AND medical_specialty = unknown AND admission_type_id = Emergency AND number_diagnoses = (5, 10] AND discharge_disposition_id = D1 → readmitted = NO

Table 5 Class association rules generated by Apriori with minconf = 0.70

No.	Rules	Confidence
1	num_procedures = (−2, 0] number_emergency = (−10, 0] number_inpatient = (−5, 0] → readmitted = NO	0.72
2	discharge_disposition_id = D1 number_inpatient = (−5, 0] time_in_hospital = (0, 5] → readmitted = NO	0.71
3	diag_1 = Diabetes num_procedures = (−2, 0] number_outpatient = (−4, 5] repaglinide = No → readmitted = NO	0.71
4	glipizide = No medical_specialty = unknown metformin = No number_outpatient = (−4, 5] race = Caucasian repaglinide = No A1Cresult = None → readmitted = <30	0.70
5	admission_source_id = AS7 admission_type_id = Emergency glipizide = No medical_specialty = unknown metformin = No number_outpatient = (−4, 5] race = Caucasian → readmitted = <30	0.70

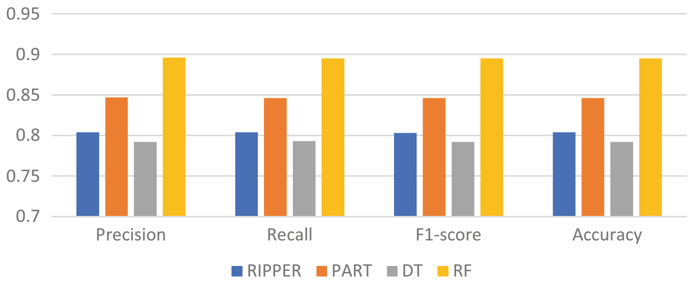


Fig. 3 Performance comparison of RIPPER, PART, decision tree, and random forest

Similarly, the first rule generated by the RIPPER algorithm suggests that if a patient has a large number of laboratory procedures and is diagnosed with a circulatory system during the stay in hospital is likely to be readmitted. Moreover, comparing the length of the generated rules by RIPPER and PART, it has been observed that PART has generated shorter and concise rules. However, RIPPER has generated fewer numbers of rules that PART. It is to be noted that for Apriori, RIPPER, and PART algorithms, only the sampled dataset based on primary diagnosis has been used.

In Fig. 3, there is a performance comparison among a regular classifier, and two rule-based classifiers have been demonstrated. The chart depicts that the rule-based classifiers have outperformed the decision tree-based classifier. Though they could not beat the random forest ensemble classifier, in terms of interpretability of the rules and overall performance, RIPPER and PART algorithms have shown good performances. Between two rule-based predictive models, the PART algorithm has outperformed the RIPPER algorithm having 84.6% accuracy, while the other one has 80.6% accuracy. The overall performance of this study is better than the other existing works too. In Table 6, a comparison with a few other existing works has been shown.

Table 6 Performance comparison with other works

Works	Test accuracy (%)
Cui et al. [6]	81.02
Ossai et al. [7]	82.97
Rubin et al. [23]	70
The proposed method (RF)	**89.5**
The proposed method (PART)	**84.6**

5 Discussion

Predicting early hospital readmission of patients can be helpful to reduce hassles to treat them properly. It can also minimize the costs associated with readmission. For diabetic patients, a decision support system that can filter out the patients susceptible to readmission is important too. This study is an effort to develop such a system with better understandability and interpretability. As per the objective of this study, classifiers have been created using various algorithms and have analyzed their performances. It has been proved that both typical and rule-based classifiers can produce good performing models. Few issues are to be discussed based on the findings from various experiments. One of the major challenges during experimentations was the high imbalance in the dataset on which has been worked on. To manage this problem, multiple measures have been taken, and the synthetic minority oversampling technique has been proven to work better than other techniques. Generating synthetic data points and adding them to the original dataset may cause the classifiers to behave differently in the real-world environment. Collecting more data for the minority class can be helpful to develop more sustainable classifiers. The experiments have been carried out to feature selection too. The empirical data shows that with feature selection, the classifiers can be robust. Among typical classifiers, the random forest algorithm has yielded better performance and has beaten currently existing works in this domain.

The main research outcome of this study lies in the rule-based predictive models. There have not been many works with those methods in this domain-specific classification tasks. Such models can help the stakeholders in a unique understandable way. By looking at the generated rules, the doctors and patients can easily grasp the way the models work. Although, the large number of rules generated by them can be overwhelming for them. Further works on reducing the rules and exploring ways to simplify them can add an extra dimension to this type of decision support system. Among the rule-based predictive models, the PART algorithm has the best performance. Though it has generated more rules than the RIPPER algorithm, the rules are short and simple. There are some limitations to this study too. To reduce the complexity of the classifiers, the three classes problem has been treated as a binary classification problem. The “>30” class has been merged with the “NO” class as interested in predicting early readmission. Although this issue could easily be resolved by conducting experiments for all these three classes, but that would add dis-ambiguity to the interpretation of the developed systems. Not utilizing more traditional algorithms is also a limitation of us. Only the decision tree and random forest algorithm have been applied. The other classification algorithms, like naive Bayes, SVM, logistic regression, etc., could be used for classification purposes, and their comparative analysis could add extra weights to the study. Despite the limitations, the outcome of this study has been good, and with the concepts presented by it, hospital readmission of diabetic patients can easily be predicted.

6 Conclusion

Ensuring better health care for risky patients even after the discharge is important. High-risk diabetic patients should be monitored carefully at the time of discharge. It should be ensured that they do not get readmitted to the hospital soon. Lacking in the follow-up treatment may increase the mortality rate and cost of health care too. To prevent these, an automated classifier that predicts the hospital readmission among diabetic patients can be useful. With that view, in this study, the unique idea of rule-based classifiers has been introduced. These classifiers can predict the 30 days readmission of a patient in an understandable way. This study has developed a few such systems, and it has been shown that with empirical data, they can surely work as a good decision support system.

References

1. Swain MJ, Kharrazi H (2015) Feasibility of 30-day hospital readmission prediction modeling based on health information exchange data. Int J Med Inform. https://doi.org/10.1016/j.ijmedinf.2015.09.003
2. Dharmarajan K, Hsieh AF, Lin Z, Bueno H, Ross JS, Horwitz LI, Barreto-Filho JA, Kim N, Bernheim SM, Suter LG, Drye EE, Krumholz HM (2013) Diagnoses and timing of 30-day readmissions after hospitalization for heart failure, acute myocardial infarction, or pneumonia. JAMA J Am Med Assoc. https://doi.org/10.1001/jama.2012.216476
3. Rubin DJ (2018) Correction to: hospital readmission of patients with diabetes. Curr Diab Rep. https://doi.org/10.1007/s11892-018-0989-1
4. Lutz R: Patients with diabetes often readmitted for hypo- and hyperglycemia. https://www.hcplive.com/view/patients-with-diabetes-often-readmitted-for-hypo-and-hyperglycemia
5. Alloghani M, Aljaaf A, Hussain A, Baker T, Mustafina J, Al-Jumeily D, Khalaf M (2019) Implementation of machine learning algorithms to create diabetic patient re-admission profiles. BMC Med Inform Decis Mak. https://doi.org/10.1186/s12911-019-0990-x
6. Cui S, Wang D, Wang Y, Yu PW, Jin Y (2018) An improved support vector machine-based diabetic readmission prediction. Comput Methods Programs Biomed. https://doi.org/10.1016/j.cmpb.2018.10.012
7. Ossai CI, Wickramasinghe N (2020) Intelligent therapeutic decision support for 30 days readmission of diabetic patients with different comorbidities. J Biomed Inform. https://doi.org/10.1016/j.jbi.2020.103486
8. Hammoudeh A, Al-Naymat G, Ghannam I, Obied N (2018) Predicting hospital readmission among diabetics using deep learning. Procedia Comput Sci. https://doi.org/https://doi.org/10.1016/j.procs.2018.10.138
9. Ribeiro MT, Singh S, Guestrin C (2016) Model-agnostic interpretability of machine learning
10. Ribeiro MT, Singh S, Guestrin C (2016) Why should I trust you? Explaining the predictions of any classifier. In: Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining. https://doi.org/https://doi.org/10.1145/2939672.2939778
11. Strack B, Deshazo JP, Gennings C, Olmo JL, Ventura S, Cios KJ, Clore JN (2014) Impact of HbA1c measurement on hospital readmission rates: analysis of 70,000 clinical database patient records. Biomed Res Int. https://doi.org/https://doi.org/10.1155/2014/781670
12. Bhuvan MS, Kumar A, Zafar A, Kishore V (2016) Identifying diabetic patients with high risk of readmission
13. Harisekaran SSDG (2015) Big data analytics predicting risk of readmissions of diabetic patients. Int J Sci Res

14. Alajmani S, Elazhary H (2019) Hospital readmission prediction using machine learning techniques. Int J Adv Comput Sci Appl 10. https://doi.org/https://doi.org/10.14569/IJACSA.2019.0100425
15. Shameer K, Johnson KW, Yahi A, Miotto R, Li LI, Ricks D, Jebakaran J, Kovatch P, Sengupta PP, Gelijns A, Moskovitz A, Darrow B, Reich DL, Kasarskis A, Tatonetti NP, Pinney S, Dudley JT (2017) Predictive modeling of hospital readmission rates using electronic medical record-wide machine learning: a case-study using mount sinai heart failure cohort. In: Pacific Symposium on Biocomputing (2017). https://doi.org/https://doi.org/10.1142/9789813207813_0027
16. Aggarwal CC (2015) Data mining. Springer International Publishing, Cham. https://doi.org/https://doi.org/10.1007/978-3-319-14142-8
17. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority oversampling technique. J Artif Intell Res. https://doi.org/10.1613/jair.953
18. Han J, Kamber M, Pei J (2012). Data mining: concepts and techniques. https://doi.org/10.1016/C2009-0-61819-5
19. Vinet L, Zhedanov A (2010) A "missing" family of classical orthogonal polynomials. Math Intell. https://doi.org/10.1088/1751-8113/44/8/085201
20. Fürnkranz J, Widmer G (1994) Incremental reduced error pruning. In: Machine learning proceedings 1994. https://doi.org/https://doi.org/10.1016/b978-1-55860-335-6.50017-9
21. Cohen WW (1995) Fast effective rule induction. In: Machine learning proceedings 1995. https://doi.org/https://doi.org/10.1016/b978-1-55860-377-6.50023-2
22. Frank E, Witten IH (1998) Generating accurate rule sets without global optimization. In: Proceedings of fifteenth international conference on machine learning. https://doi.org/1-55860-556-8
23. Rubin DJ, Handorf EA, Golden SH, Nelson DB, McDonnell ME, Zhao H (2016) Developement and validation of a novel tool to predict hospital readmission resk among patients with diabetes. Endocr Pract. https://doi.org/10.4158/E161391.OR

Design and Analysis of Mobile-Based Tourism Security Application: Concepts, Artifacts and Challenges

Deepanjal Shrestha, Tan Wenan, Bikram Adhikari, Deepmala Shrestha, Adesh Khadka, and Seung Ryul Jeong

Abstract Tourism security has become a matter of great concern over the years for tourists around the world, and this attribute plays an important role in the selection of a destination. The application of technology can greatly enhance the safety and security environment of a tourist in a particular destination. This work proposes the analysis and design of a mobile-based application targeted to provide security and safety to a tourist. The design is based on requirements gathered from the tourist and tourism business personals through online and offline interviews which are grouped and coded into user requirements. Unified Modeling Language (UML) basic notation and tools are used to analyze and design the system in a comprehensive manner. Pokhara city of Nepal is chosen as the destination to deploy destination city to deploy and test the system. Further, the challenges and limitations of the work are discussed in light of privacy and system implementations. This work has a great significance in the

---

D. Shrestha · T. Wenan
School of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China
e-mail: deepanjal@hotmail.com

T. Wenan
e-mail: wtan@foxmail.com

T. Wenan
School of Computer and Information Engineering, Shanghai Polytechnic University, Shanghai, China

B. Adhikari
Genese Solution, Kathmandu, Nepal
e-mail: bikram@genesesolution.com

D. Shrestha · D. Shrestha
School of Business, Pokhara University, Pokhara, Nepal
e-mail: deepmala@pusob.edu.np

A. Khadka
Ministry of Education, Science and Technology, Singhadarbar, Kathmandu, Nepal
e-mail: adesh.khadka@gmail.com

S. R. Jeong (✉)
Graduate School of Business IT, Kookmin University, Seoul, South Korea
e-mail: srjeong@kookmin.ac.kr

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_15

application of technology for personal security especially in the tourism area. It is one of the first kind of work done in the tourism sector of Nepal which adds a great value to the tourism industry. The work also contributes as a knowledge domain for the digital and mobile technology implementations in tourism destinations.

Keywords Unified Modeling Language (UML) · Global Positioning System (GPS)

1 Introduction

Tourism business around the world has accounted for 10.4% of global GDP and created 319 million jobs as per the report of World Travel and Tourism Council in 2018. The report also states domestic tourism had the strongest growth with 71.2% of the total spending in 2018 [1]. A lot of factors account for the growth of tourism business around the world which includes tourism environment, policies, infrastructure, cultural and natural resources [1], safety and security, and health and hygiene. Tourism safety has become a greater concern for the tourist around the world, and a lot of attention is paid to travel safety by the tourist. Tourism safety can be managed and handled properly with the use of technology. The IoT devices, social sites, Global Positioning System (GPS) technologies, image satellites, CCTVS, etc., are very powerful technological tools that are capable of working with security and disaster-related situations [1]. Mobile devices are the handiest and preferred devices that can be used as a personnel means of communication to get connected with individuals during sensitive situations. Mobile in tourism security is the best tool that can provide easy communication and act as an effective medium to avoid disaster and risk. The role of mobile devices and applications can be extremely vital for underdeveloped countries like Nepal which are in the middle of the technological boom and has concerns over its tourism business. Further, the data from Nepal Tourism Ministry depicts that in 2018, 619 cases related to tourism safety were reported which mostly include cases related to personal incidents. Cases with women, aged tourists, and minors were also seen in existence. Data published in 2017 by The Himalayan Times reported that 62 tourists have gone missing in the last five years which was alarmingly big for safe tourist destination image [2]. In light of these incidents, it can be inferred that the tourism security system is a must and one of the most important systems required. Thus, this research work is undertaken to analyze, design, and develop a mobile application for tourism safety in Nepal. The requirements are gathered from tourists, tourism stakeholders, and associated literature in an iterative cycle to come up with a mobile-based software requirement specification document. UML tools and techniques are used to design and develop the system. Easy to use interface with a priority of action and sequences during risky situations are taken as major considerations. The application design is of hybrid nature that works with mobile devices, tablets, PCs and computer systems having Internet access. The system is targeted for Pokhara city of Nepal, which is the tourist capital and the second-biggest city of Nepal.

2 Literature Review

Tourism security and technological implementations have taken the attention of a wide range of researchers around the world. Many studies have been done in the past that talk about design issues, social impacts, business impacts, future trends AI role in tourism security, and disaster management. Andriy Volkoviy and Llya Tikhov in their paper, Use of mobiles in security and safety systems explored all the different possibilities that can be implemented with mobile to provide safety to the users. They listed some good mobile applications like, Rescue, iMap Weather Radio, Earthquake Alert, Real-Time Warning, and Tsunami in the paper with their features and shortcomings [3]. Similarly, studies have been done on safety forecasting and early warning [4] safety of food services in the hospitality industry [5] safety and security in different geographies [6] and mobile application as guides integrated with security features [7] and using social site data in understanding tourist pattern for better management [8]. The book written by Michael and Annika extensively talks about the use of mobile applications frequently on vacation, where they talk about the varied uses of mobile applications in tourism destinations [9]. Studies are also carried on the role of mobile applications in tourism [10] and mobile application in women safety [11–14].

Table 1 represents a survey of some recent research work in the field of tourism, women's safety, and mobile technology. Further, the literature work regarding tourism risk and safety of Nepal has portrayed Nepal as a risky country with 84th rank out of 176 countries [1]. Various studies have been conducted on disaster management, risk mitigation, and adventure tourism in Nepal but very less data is available for technology and tourism. The data of Table 2 depicts the technological development of Nepal with technological breakthroughs (mobile, internet, and social site) and the current status quo [15]. The technological development has seen few published works of few notable authors Shrestha and Jeong, Bidur et al. Tan et al. [16] who talk about the role of technology in tourism but there is a gap and absence of literature in the case of mobile technology and tourism security.

3 Research Framework

The research framework plays a vital role in guiding the overall study of the research work. In our study, UML is used as a basic tool for the analysis and design of a mobile application for the tourism industry. Figure 1 depicts, five classes of the user (tourist, tourism business personnel, tourist officials, security agencies) who provide input to the system requirement. The requirements are gathered using an iterative approach which results in identifying concepts, use cases and systems artifacts which are further elaborated to come up with the development of the conceptual model, use case diagram, and UI model for the system. Some specific algorithms of tourism security are also proposed based on user requirements in Nepal as shown in Fig. 1.

Table 1 Survey of mobile technology and application in tourism safety

Paper details (year/publisher)	Research details	References
2012, Journal of Information and Security	Use of mobiles in security and safety systems	[3]
2015, IEEE Digital Explore	An android app for the safety of women	[11]
2015, Journal of Computer Science and IT	Android application for women security	[13]
2015, Journal of Computer Applications,	Mobile-based intelligent safety system for women	[14]
2015, KSII Transactions	Exploring the attractive factors of app icons	[17]
2017, KSII Transactions	An app visualization design for car accident prevention	[18]
2018, CISTI Conference	Technology usage for safety and security	[6]
2019, ICCPCT (Conference)	Mobile application for women safety and RT database	[12]
2019, Journal of Physics	Mobile apps in tourism communication	[10]
2019, Third World Conference	Twitter mining sentiment analysis in tourism industry	[8]
2019, IEEE Access	Safety forecasting and early warning	[4]

4 System Analysis and Design (Concepts and Artifacts)

The section below represents different concepts, artifacts, and processes identified and used for the development of mobile-based security system application.

4.1 User Requirements and Functional Description

User interview and narrative description of users which include tourist, tourism business personnel, tourist officials, and security agencies are recorded. The requirements are analyzed to identity artifacts, objects, processes, and functions. The functional requirements of the system are shown in Table 3. The functional requirements are coded and categorized as evident and hidden depending on their state in the system. Altogether ten major categories of functional requirements are identified numbered from R0 to R9, with each major category further extended to specific functional requirements. The extended functional requirements are coded and grouped based on major functional types, also identified as evident and hidden.

Table 2 Data representing technological development in Nepal [2, 15]

ICT Development	Year	Details
E-readiness and digital implementations	2016	135 as per world bank
Government development index rank	2017	1.85 as per maturity
Mobile communications	2013–2022 (forecasted)	25–136% (forecasted)
Mobile Internet penetration	2012–2017	21–57% by 2022
Internet penetration	2012–2017	21–57%
Active Internet users	2018	250 new Internet users per hour
Social site users	January 2018	9.3 M Facebook, 604 YouTube
Web sites status	2017	56,286 registered Web sites in Nepal
4G subscribers	2017	2,144,887
Broadband services	2020	90% of population gets connected
E-commerce	2018	Recorded a growth rate of 62% CAGR
Online travel	2018	Recorded a growth rate of 15% CAGR
Online media	2018	Recorded a growth rate of 44% CAGR

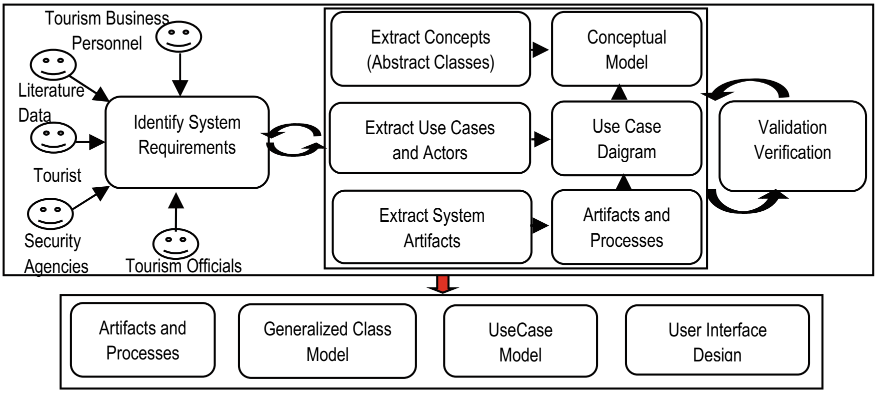


Fig. 1 Research framework of the study with detailed components

Table 3 User requirements for mobile security application

Reference	Function description	Category
R0	User registration requirements	Evident
R0.1	Get information details of the user who wants to use this application	Evident
R0.2	Check information with the database to find out errors, redundancy, format, etc.	Hidden
R0.3	Register information of the user after validation and provide a unique ID	Evident
R1	Destination-based information requirements	Evident
R1.1	Legal issues in sensitive places like airport for drugs, cash, gold, banned items, etc.	Evident
R1.2	Administrative issues: Visa regulation, fee	Evident
R2	Warning and alert message requirements	Evident
R2.1	On travel for taxi fare, bus fare, or other travel vehicles	Evident
R2.2	Sensitive situations like accidents, disaster, road blocks, rallies, etc.	Hidden
R3	Information requirements	Evident
R3.1	Alerts on destinations with regard to health and hygiene	Hidden
R3.2	Alerts on routes, services, prices, support, digital communication	Evident
R3.3	Information on special regulations for religious, cultural, sensitive places	Evident
R3.4	For change in tourism destinations, activities, and critical updates	Hidden
R4	Connection and information sharing requirements	Evident
R4.1	Police services, rescue agencies, local governance, etc.	Evident
R4.2	Tourism information centers and agencies	Evident
R4.3	Emergency rescue teams and agencies	Evident
R4.4	Nearby tourists or peers in existence	Evident
R5	Risk sensing and mitigation requirements	Evident
R5.1	Sense any incoming risk situation and provide solutions	Hidden
R5.2	Track the location and provide rescue	Hidden
R5.3	Provide backup plans for any risk situation	Evident
R6	Personal safety requirements	Evident
R6.1	Track and individual to ensure personal safety and inform for suspicion	Evident
R6.2	Provide safety to personal data including data and information	Hidden
R6.3	Prioritize solo females, aged, or very young tourist	Evident
R7	User application requirements	Evident
R7.1	Provide easy interface to trigger emergency situation	Evident
R7.2	Provide real-time sensing and monitoring feature	Hidden
R7.3	Provide information on safe routes or alternative routes short routes	Evident
R7.4	Provide information on safe food and hygiene locations	Evident
R7.5	Provide analysis of the risk of a visiting a place and provide instructions	Evident
R7.6	Provide features to report a complain	Evident

(continued)

Table 3 (continued)

Reference	Function description	Category
R7.7	Provide features for historical data of incidents and mishappenings	Evident
R7.8	Provide alarms, alert messages for any risk or disaster information update	Hidden
R7.9	Provide information recording and location-sensing feature for emergency	Hidden

4.2 Conceptual Design

Concepts define the entities of the system at an abstract level that can be later refined to form software class and objects which interact with one another to provide a system functionality [19].

4.2.1 Concept Extraction Process

The process of use case and concept extraction is subject to user requirements which are traced from interview data. For example user requirement R0 that states that a user must be a registered user to use the security application. The user gets registered by providing the necessary data and information as requested by the application. In this case, required data here is a user name, country, mobile, or passport number, whereas the process is to validate and register users by the system. This requirement is related with a major use case Register_User. Further, Register_User has further dependencies to complete this process which include get info, check info, process registration, and invoke OTP. The conceptual class person (tourist/general user) invokes this use case and provides all the necessary information and gets himself stored as a user object in the system. Further in Register_User, use case process needs other conceptual classes like policy information from governing bodies, communication roles from devices classes, and other objects to validate and complete this process. Thus, user requirements and use cases serve as the base to extract the concepts. A similar process when mapping requirements can be applied to design concepts and map them as conceptual models in the early requirement and detailed design phase. Figure 2 shows the conceptual class modeled in a generalized relationship with few extended as inherited classes.

4.2.2 Generalized Conceptual Model

This section depicts the process of identification of concepts in a tabular form extracted from user requirements and other design artifacts of the system. Table 4 depicts a class person that extends the types of users which include tourist type (inbound, domestic), tourism personnel, and government officials. All identified

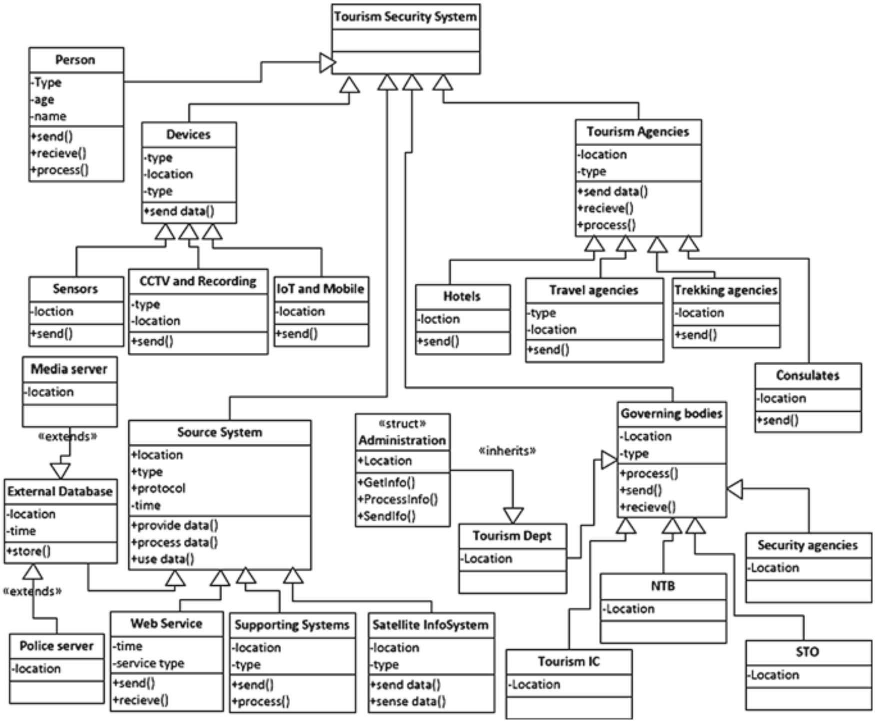


Fig. 2 Conceptual model of security system is depicted with a generalized relationship

Table 4 Representing base class person with other attributes and extended classes

Base class	Extended classes	Properties	Functions	Multiplicity
Person (Users)	Local citizen	Name, age, address	Post data, use data, get information	*….1
	Domestic tourist	Name, age, address	Post data, use data, get information	*….1
	Inbound tourist	Name, age, address	Post data, use data, get information	*….1
	Tourism personnel's (Media, tourism labor)	Name, age, address	Post data, use data, get information	*….1
	Government officials	Name, age, address	Post data, use data, get information	*….1

concepts have properties, functionalities, relationships, and multiplicities associated with them. Similarly, the generalized systems represent other important base concept classes which include the source systems, governing bodies, tourism agencies, and devices class at a higher level. Each base concept class is further extended into many generalized classes that further includes the specialized classes needed by the system to form the real artifacts of the system. The objects of the classes form the dynamic part of the system that is refined from the conceptual model and use the case model in the later phase of the design process in an iterative manner. The generalized conceptual model is shown in Fig. 2 that represents the conceptual classes of the mobile-based security system.

Further, the class concept of external database and source system are important concepts that are related to information provider and consumer objects which include media servers, police servers, etc. The device class in the diagram is a dedicated information providers related to Sense_Data; use case and includes sensors, mobile devices like smartphones, iPad, CCTVS, recording system, and devices. These devices are capable of performing functions that include sending data, receive instructions, post data, and can create information instances. The class tourism agencies act as an important actor during rescue operations so it is taken as a default security agency for the tourist and is related to many use cases which include Create_Circle, Invoke_Panic, and Connect_Agencies. The embassies and consulates are more concerned with the well-being of their citizens. Similarly, the class governing body is represented with associated classes Department of Tourism, Security Agencies, Nepal Tourism Board (NTB), State Tourism Office (STO), and Tourism Information Centers (TIC). These classes are related to security policymaking, management, and implementation of it.

4.3 Use Case Model

Use cases are the stories that illustrate and imply the requirements of a system in an informal way [19]. Use cases are the narrative descriptions that help in identifying the system requirements that can be further elaborated to form a sequence diagram, class diagram, and collaboration diagram representing the static and dynamic view of the system. The section below represents various high level and extended use cases of the system.

4.3.1 The High-Level Use Cases

The high-level use cases are created at the initial level of system analysis and design. It is more abstract and gives a good understanding of the system processes that include one or many functions to complete a particular task of the system [19]. Table 5 represents 13 major uses cases of the mobile security application where each major use case is related to the number of extended use cases that capture the complete

Table 5 Representing major tourist (user) initiated use cases and extended use cases

Major use cases	Extended use cases	Actors
Register_User	Get info, check info, process registration, invoke OTP	Tourist, user
Create_Circle	Get contact list, add contact, add message, define level	Tourist, user
Emergency_Rescue	Send info, display security agencies, contact rescue agency, monitor real time	Tourist, user
Real-time_Monitoring	Capture location, sample info, check status, issue alerts, mark tourist	Tourist, user
Invoke_Assistance	Ask Info, check database, compute info, provide info	Tourist, user
Plan_Trip	Collect travel info, check database, create plan, display plan, provide instructions	Tourist, user
Route_Navigate	Get location, compute alternate paths, get related data, provide suggestion	Tourist, user
Security_Update	Check database, capture changes, display info	Tourist, user
Share_Location	Capture location, send data, send message	Tourist, user
Connect_Peers	Inform peers, inform circle, display peers, monitor real time	Tourist, user
File_Complaint	Open form, post data, upload info, submit data	Tourist, user
Offline_Assistance	Invoke backup, search request, provide info	Tourist, user
Invoke_Panic	Get location, alter agencies, make rescue, update info	Tourist, user

functionality of the system. As seen, the major use case Create_Profile requires Get_info, Check_info, Process_Registration, and Invoke_OTP use cases to complete its functional process. A person is an actor that initiates this process and closes it. Similarly, we can relate, understand and deduce other uses cases of the system domain. The actors for these use cases are a person that represents domestic, inbound, and tourism personnel. The other uses cases include Create_Circle, Emergency_ Rescue, RT_ Monitoring, Invoke_ Assistance, and Plan_Trip, etc., as shown in Table 5.

4.3.2 Use Case Diagram

Use case diagrams are the pictorial representation of the use cases that show the system processes in a diagrammatic form [19]. Figure 3 represents complete system use cases including their actors and system boundary. The use case diagram represents 13 major use cases that are triggered by tourist and user and seven major use cases that are triggered by the external server system. The tourist use cases are connected with black lines, while the users with red lines and systems with blue lines as shown below. The mobile application, in this case, serves as the system boundary and is a part of a larger system, the security management system. Further, in Table 6, the sequence of events is explained between the actor and system to show the implementation of various use cases in the system. The sequence of events can

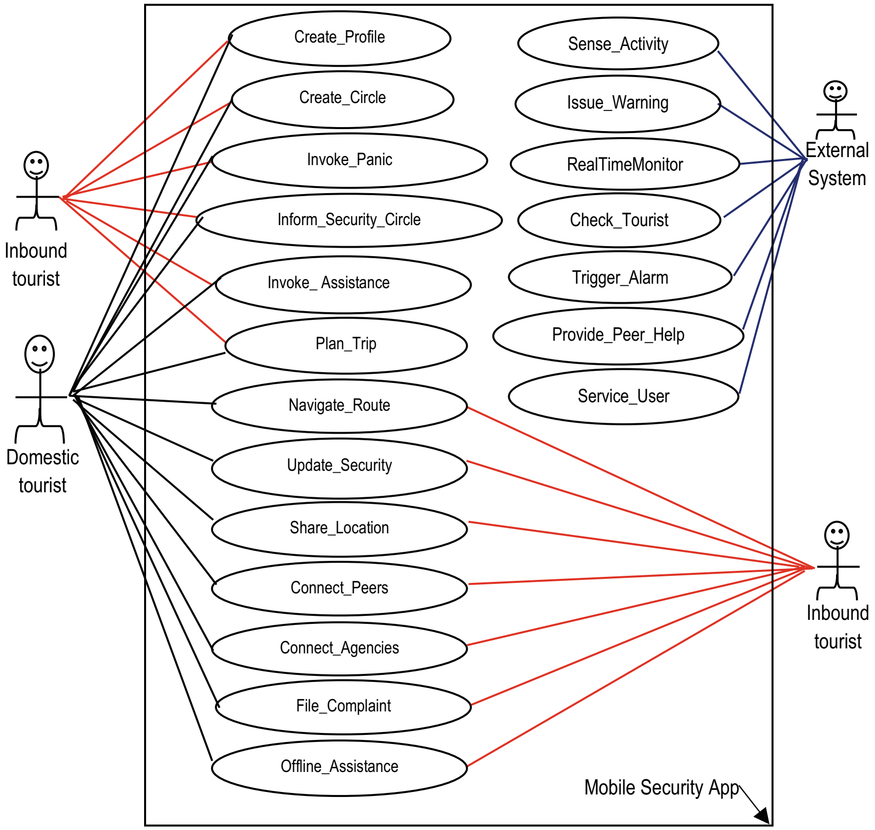


Fig. 3 Use case diagram of the mobile security system application with related actors

Table 6 Typical course of events for use case: Invoke_Application

Actor action		System response	
1	User initializes application with a touch	2	System prompts for user input
3	User inputs preferred details	4	System validates input types (options)
5		4a	Logs in as register user and waits for the next action
6	User acts with information	4b	Logs in as guest user and waits for the next action
		4c	Initiates panic module and distribute information of the user
Dependency		**GetUser_info, Check_Input_Type, Validate_Input_Type, Trigger_Action**	

later be extended to come up with sequence diagrams and collaboration diagrams that show the behavioral design of the system. Table 6 shows a typical course of events for Invoke_Application that has three optional paths that are shown as 4a., 4b. 4c in course of events. Table 6 also shows the dependency of these primary use cases on other use cases which include GetUser_info, Check_Input_Type, Validate_Input_Type, and Trig-ger_Action to complete a particular system process.

4.4 The Radius R-algorithm

The radius R algorithm is a security recommender system that works for providing travel route and blockade/accident scenario checkup suggestions to the tourist-based on their current location and destinations as shown in Fig. 4. The algorithm is specially designed for the security scenario of Nepal considering its important parameters. The system is named as radius R algorithm because it takes radius as its main parameter to compute paths and trace agencies for making recommendations from the specified or current node. In case, the tourist meets an accident at node (n), then also system considers radius R to make a scan to locate health agencies, security agencies, and rescue teams so that appropriate recommendations can be made. The radius is set to 2 km in case of metropolitan and sub-metropolitan areas while the radius is set to 4 km in case of rural areas because the metro and sub-metro are stuffed with health, security, and rescue facilities closely which is not present in the case of rural areas. On the first scan, if the defined radius area finds no health, security, or rescue agency, the algorithm increases the radius by 50% and makes the next cycle of scan again, repeating further.

This process keeps increasing the radius by 50% each time until agencies or nodes are spotted. The same logic applies to recommend routes and paths also. Besides this,

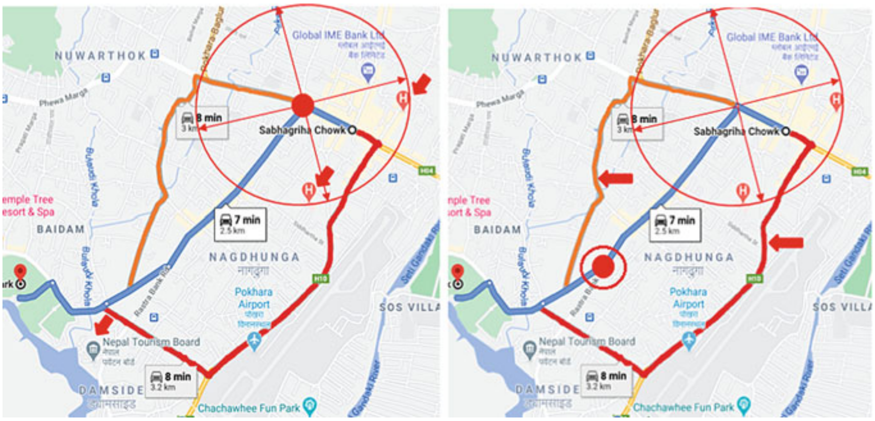
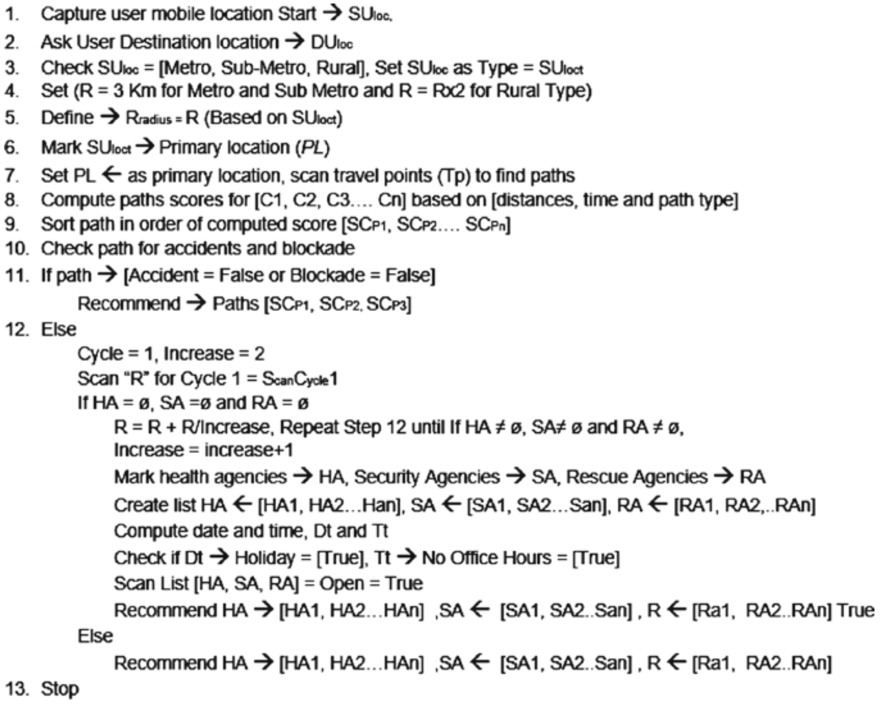


Fig. 4 Representing different conditions for route finding and recommendations

the algorithm synchronizes its data for day and time with the stored calendar of Nepal to see the holiday period and working hours of a concerned agency which is further computed to make accurate recommendations as explained in Fig. 5.

Table 7 shows the comparison of attributes of R algorithm with some of the other algorithms, and it is seen that this algorithm has some features that are not present in other algorithms like the backtracking of the path and recommending agencies based on radius. Further, this algorithm is improvised to get synchronized with the Nepal government officials calendar data to make realistic recommendations. The R algorithm is applicable for smart transportation, smart city management, security, and disaster management applications as it combines the features of searching the travel paths, travel points, and security agencies located in the circular periphery with a backtracking system. The adjustment of an algorithm to keep increasing the radius with scan cycles to locate a valid information cluster is a strong feature. The other algorithms that were specifically targeted for the Nepalese tourism security system included the connect peers and rescue algorithm along with other general algorithms.



```
1.  Capture user mobile location Start → $SU_{loc.}$
2.  Ask User Destination location → $DU_{loc}$
3.  Check $SU_{loc}$ = [Metro, Sub-Metro, Rural], Set $SU_{loc}$ as Type = $SU_{loct}$
4.  Set (R = 3 Km for Metro and Sub Metro and R = Rx2 for Rural Type)
5.  Define → $R_{radius}$ = R (Based on $SU_{loct}$)
6.  Mark $SU_{loct}$ → Primary location (PL)
7.  Set PL ← as primary location, scan travel points (Tp) to find paths
8.  Compute paths scores for [C1, C2, C3…. Cn] based on [distances, time and path type]
9.  Sort path in order of computed score [$SC_{P1}$, $SC_{P2}$…. $SC_{Pn}$]
10. Check path for accidents and blockade
11. If path → [Accident = False or Blockade = False]
        Recommend → Paths [$SC_{P1}$, $SC_{P2}$, $SC_{P3}$]
12. Else
        Cycle = 1, Increase = 2
        Scan "R" for Cycle 1 = $S_{canCycle}$1
        If HA = ø, SA =ø and RA = ø
            R = R + R/Increase, Repeat Step 12 until If HA ≠ ø, SA≠ ø and RA ≠ ø,
            Increase = increase+1
            Mark health agencies → HA, Security Agencies → SA, Rescue Agencies → RA
            Create list HA ← [HA1, HA2…Han], SA ← [SA1, SA2…San], RA ← [RA1, RA2,..RAn]
            Compute date and time, Dt and Tt
            Check if Dt → Holiday = [True], Tt → No Office Hours = [True]
            Scan List [HA, SA, RA] = Open = True
            Recommend HA → [HA1, HA2…HAn] ,SA ← [SA1, SA2..San] , R ← [Ra1, RA2..RAn] True
        Else
            Recommend HA → [HA1, HA2…HAn] ,SA ← [SA1, SA2..San] , R ← [Ra1, RA2..RAn]
13. Stop
```

Fig. 5 Radius "R" algorithm structure

Table 7 Comparison of R algorithm with some existing systems and algorithms

Attributes	Captures location of the user	Suggest shortest path and time	Suggest places based on sensing (Mishappening, roadblocks)	Suggest location based on circular area (radius) and backtracking	Works with Nepal government data for holidays, office hours for recommendations	References
R algorithm	Yes	Yes	Yes	Yes	Yes	–
safety forecasting and early warning	Yes	Yes	Yes	No	No	[4]
Intelligent safety system	Yes	Yes	Yes	No	No	[8]
Distributed societal security system	Yes	No	Yes	No	No	[11]
Google map	Yes	Yes	No	No	Yes	[20]
Dijkstra’s algorithm	No	Yes	No	No	No	[20]
Any angle path planning	No	Yes	No	No	No	[20]
Mobile/IoT-based Algo	Yes	Yes	Yes	No	No	[21]
Mobile in security and safety	Yes	Yes	Yes	No	No	[22]

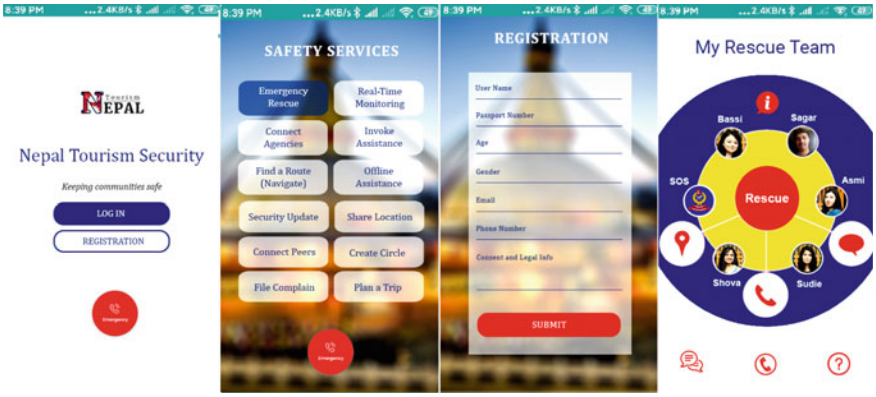


Fig. 6 Snapshots of screens including welcome, services, rescue, and registration

4.5 User Interface Design

The user interface design serves as the heart of the whole system process, and several designs were reviewed before coming up with the interface of the application. The more sensitive requirements were the given priority in user interface design and icons. The GUI principles especially for visually targeted web applications were considered with more attention to color and text size. Few snapshots of the user interface are depicted in Fig. 6.

5 System Design Verification and Validation

The system analysis and design are verified and validated using design walkthrough and backtrack tracing in the case of different models developed [19]. It can be seen from Table 8 that class user/tourist is related to use case Create_Profile which has tourist/user as an actor, and it can be traced back to user requirement R0, which consists of sub-functions R0.1, R0.2, and R0.3, respectively. Similarly, class source systems can be related to use case Realtime_Monitor, Update_Security, which is further related to actor tourist/user and server system and can be traced back to requirements R1, R2, R3, and R4, respectively. This shows that analysis and design are in coherence with the user requirements and follow a connection between design artifacts to guarantee a good design.

Table 8 Verification and validation matrix

Class	Use case	Actor	Traced to requirement
User/tourist	Create profile	Tourist/user	R0
Source system	Realtime_Monitor, Update_Security, Navigate_route, Sense_Activity	Tourist/user, Server system	R1, R2, R3, R4
Device systems	Realtime_Monitor, Update_Security, Navigate_route, Sense_Activity	Tourist/user, server system	R1, R2, R3, R4
User/tourist	Plan_Trip, Navigate_Route	Tourist/user, server system	R1, R3, R6, R7
User/tourist	Plan_Trip, Navigate_Route	Tourist/user, server system	R1, R3, R6, R7

6 System Security and Testing

Digital systems and mobile systems are prone to attack from hackers, malicious bots, viruses' worms, Trojan horses, etc. To provide security to the application, current security measures are used in the system that includes an SSL certificate, strong password policy, mobile or web-based OTP, and registration based on either passport number or mobile phone to authenticate. The guest login is limited to view the only privilege. The password is assigned 15 days of life, and the user has to go with a two-factor authentication after the period expires. To protect the personal data of users, the database is designed to show no policy and is centrally stored in a secure server in a two-layered system that can only be accessed by authorized government officials through verification and validation of their identity. The testing of the system is done with user-generated data and in a controlled environment. Almost all the modules and paths are tested in the application and data results are satisfactory. The work lacks testing with real-world data and real-world implementation. Performance, load balancing, and recommendation preciseness need to be tested together for 100% accuracy.

7 Challenges and Limitations of the System

The design and implementation of the mobile-based tourism security application have different challenges associated with it. Infrastructure is one of the biggest challenges for the robust implementation of the system. The security application is dependent on different sources and systems to extract data related to the security environment. The real-time collection, processing, and distribution of data require a huge digital and communication infrastructure that must be capable of processing complex data and algorithms. User data privacy and security are another major concern, as the data of individuals have a high chance of being misused, and it needs a strong legal and system policy to ensure user data safety. The optimal use of the application by users as a trusted system can be difficult due to the privacy concerns and nature of the tourist as they are from different countries, backgrounds and have issues with language, icons, and information models used. In the context of Nepal, IT education, system support, and governing policies can add more challenges to the system implementation. The designed application is good for tourism recommendation with security as the leading priority but still suffers from issues of cyber-attacks and misuse of personal data from officials having access to it. It is also dependent on external sources for data and information.

8 Conclusion

The above study concludes that security applications are one of the most vital needs for the tourism industry to guarantee safety as well as create image of the destination. A secured tourism destination has more chances of getting tourists from all sections and age groups compared to unsafe destinations. Technology can be aggressively used to implement security in the tourism sector at various levels. An integrated tourism security system with mobile application extension for smartphones can enable safety in the hands of tourists visiting a place. A good and well-analyzed mobile application can not only work as a good security system for tourist but can be enhanced as a personal security system too. Though privacy and security of personal data remains a big question for the implementation aspect, the overall realized advantage minimizes the privacy concern. A well-thought plan with stiff regulations must be imposed before implementing this kind of system. Further the design of software systems using UML tools and techniques greatly helped in identifying the base components, processes, and functions of the system. The UML approach made an analysis and design process more effective with designated notations and symbols. The designs carried with extensive research and GUI principles made it look more appealing and attractive. User-provided excellent feedback in the testing phase for the UI design, which was considered as a merit achievement. The specifically designed algorithms serves the local customized needs of the tourist in Nepal. It is concluded that the system is well designed with engineering principles, in time and has been well documented. This work serves as a good source of knowledge and reference for future systems.

Acknowledgements We would like to thank the Ministry of Tourism, Culture and Civil Aviation of Nepal for providing necessary support and data for this work. Further, this work is supported in part by the National Natural Science Foundation of China under Grants No. 61672022 and No. U1904186.

References

1. WTTC, Travel and Tourism Economic Impact 2019, World Travel & Tourism, 65 Southwark Street, London SE10HR, United Kingdom. https://www.slovenia.info/uploads/dokumenti/raziskave/raziskave/world2019.pdf
2. Aryal NP et al (2019) Nepal tourism statistics 2018. Ministry of Culture, Tourism and Civil Aviation, Kathmandu, Nepal. https://tourism.gov.np/files/statistics/19.pdf
3. Volkoviy A, Tikhov L (2012) Use of mobiles in security and safety systems, information & security. Int J 28(1):146–153
4. Yin J, Bi Y, Zheng X, Tsaur R (2019) Safety forecasting and early warning of highly aggregated tourist crowds in China. IEEE Access 7:119026–119040. https://doi.org/10.1109/ACCESS.2019.2936245
5. DusenkoSV (2018) Digital technology in ensuring the safety of food services in the hospitality ındustry,2018. IEEE ınternational conference (IT&QM&IS), St. Petersburg, pp 618–619, doi: https://doi.org/10.1109/ITMQIS.2018.8524926
6. Au-Yong-Oliveira M, Moreira F, Martins J, Branco F, Goncalves R (2018) Technology usage as a way to increase safety and security in different geographies: testimonials on the use of technology in Rio de Janeiro, Brazil,2018. In: 13th Iberian conference on ınformation systems and technologies (CISTI), Caceres, 2018, pp 1–7. doi: https://doi.org/10.23919/CISTI.2018.8399266
7. Alrehili M, et al (2018) Tourism mobile application to guide madinah visitors. In: 2018 1st ınternational conference on computer applications & ınformation security (ICCAIS), Riyadh, pp 1–4. doi: https://doi.org/10.1109/CAIS.2018.8442023
8. Gupta G, Gupta P (2019) Twitter mining for sentiment analysis in tourism ındustry. In: 2019 third world conference on smart trends in systems security (WorldS4), UK, pp 302–306. doi: https://doi.org/10.1109/WorldS4.2019.8903940
9. Beier M, Aebli A (2016) Who uses mobile apps frequently on vacation? Evidence from Tourism in Switzerland. ICT in Tourism 2016. Springer. https://doi.org/https://doi.org/10.1007/978-3-319-28231-2_40
10. Abdul R et al (2019) Mobile apps in tourism communication: the strengths and weaknesses on tourism trips. In: The 2nd joint ınternational conference on emerging computing technology and sports 2019, Bandung, Indonesia. https://doi.org/https://doi.org/10.1088/1742-6596/1529/4/042056
11. Yarabothu RS, BramarambikaThota A. An android app for the safety of women. In: 12th IEEE India ınternational conference E3-C3, At: JamiaMilliaIslamia, New Delhi, Indıa. https://doi.org/10.1109/INDICON.2015.7443652
12. Prashanth DS, Patel G, Bharathi B (2017) Research and development of a mobile based women safety application with real-time. Int Conf ICCPCT 2017:1–5. https://doi.org/10.1109/ICCPCT.2017.8074261
13. Harini R, Hemashree P (2019) Android application for women security. J Comput Sci Inform Technol 8(10):54–59. ISSN 2320-088X
14. Paradkar A, Sharma D (2015) All in one ıntelligent safety system for women security. Int J Comput Appl 130(11):33–40
15. Frost & Sullivan and Government of Nepal, Ministry of Communication & Information Technology, 2018 Digital Framework Nepal. https://mocit.gov.np/application/resources/admin/uploads/source/EConsultation/Final%20Book.pdf
16. Tan W, Shrestha D, Jeong SR (2019) Digital tourism development and sustainability model for Nepal. In: 2019 IEEE CSCWD, Portugal, 2019, pp 182–187. doi: https://doi.org/10.1109/CSCWD.2019.8791852
17. Ho C, Hou K (2015) Exploring the attractive factors of app icons. KSII Trans Internet Inf Syst 9(6):2251–2270. https://doi.org/10.3837/tiis.2015.06.016
18. Jeong Y, Jeong E, Lee B (2017) An app visualization design based on ıot self-diagnosis micro control unit for car accident prevention. KSII Trans Internet InfSyst 11(2). doi: https://doi.org/10.3837/tiis.2017.02.020

19. Larman C (2005) Applying UML and patterns, 3rd edn. Pearson Education. ISBN-13: 978-0131489066, ISBN-10: 0131489062
20. Wikimedia Foundation, Inc., Pathfinding, Creative Commons Attribution-ShareAlike License July 2020. https://en.wikipedia.org/wiki/Pathfinding
21. Bin C et al (2019) A travel route recommendation system based on smart phones and IoT environment. Wirel Commun Mob Comput. https://doi.org/10.1155/2019/7038259
22. Volkoviy A, Tikhov I (2012) Use of mobile application in security and safety systems. Inf Secur Int J. https://doi.org/10.11610/isij.2812

An Efficient Approach to Software Fault Prediction

Md. Anuvob Pradan, Mahia Binte Mizan, Moon Howlader,
and Shamim Ripon

Abstract The use of machine learning concepts in the software engineering field is now ubiquitous to predict software defects. Most of the Software Defect Prediction (SDP) datasets are highly imbalanced and filled with multiple irrelevant features that cause negative effects on the results. The goal of this paper is to create an approach to predict the software faults efficiently from imbalanced and multi-featured SDP datasets. Two highly imbalanced and multi-featured NASA MDP datasets have been used in this experiment. Initially, data cleaning has been performed with the help of *Z*-score technique to eliminate noisy instances. To balance the datasets, Synthetic Minority Oversampling Technique (SMOTE) oversampling technique has been used. Furthermore, to select the relevant features, three well-known feature selection techniques, as well as their ensembles, have been applied. Finally, to measure the performance, four well-known classification algorithms are implemented and evaluated the results by their accuracy, TPR and TNR.

Keywords Synthetic Minority Oversampling Technique (SMOTE) · Software Defect Prediction (SDP) · Machine Learning (ML) · K-Nearest Neighbor (KNN) · Logistic regression · Decision tree · Random forest

1 Introduction

Softwares are now vital for both technical and scientific purposes. As the need for good quality software increases, so does the necessity for quality assurance activities

---

Md. A. Pradan (✉) · M. B. Mizan · M. Howlader · S. Ripon
Department of Computer Science and Engineering, East West University, Dhaka, Bangladesh
e-mail: mdanuvobprodan@gmail.com

M. B. Mizan
e-mail: m.mizan0129@gmail.com

M. Howlader
e-mail: ahmoon67@gmail.com

S. Ripon
e-mail: dshr@ewubd.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_16

such as testing, verification and validation, fault tolerance, and fault prediction. Fault prediction algorithms are often used by organizations with a low budget and time constraints for detecting fault-prone modules. To improve the detection rate and performance of software fault detection, Machine Learning (ML) is considered to be the most popular method [1]. However, applying ML raises the concern of imbalanced datasets. As the Software Defect Prediction (SDP) datasets are highly skewed and filled with multiple irrelevant features, it results in biased trained models, which return faulty modules as non-faulty. This was shown in earlier works, e.g., Mohsin et al. [2]. When these models are used on various other software the cumulative effect of low true positive rate causes numerous errors to be generated.

In this paper, a novel approach to predict software faults efficiently is proposed. The proposed approach consists of a data cleaning method to reduce noisy instances, SMOTE oversampling technique to generate pseudo instances, and four feature selection algorithms, for dimensionality reduction, have been included.

In an imbalanced dataset, the results get biased toward the majority class; this, in turn, produces models with poor predictive performance, specifically for the minority class. For datasets, such as the SDP dataset where defect cases are the minority and less likely to happen than non-defect cases, the minority class is more crucial. Therefore, classes with fewer instances are more prone to classification errors than those with a higher number of instances. Oversampling techniques try to solve this problem by adding instances to the minority class either by duplicating or adding fake data. It can further be rectified by performing data cleaning first, to remove noisy instances or outliers. An outlier is a data point that differs so much from other values that it is suspected of being obtained by a separate mechanism [3].

The SDP datasets contain a large number of features, but every feature in the dataset does not influence the classification results. The initial dataset may contain irrelevant, duplicate, and useless data that do not affect the learning outcome; instead, these attributes can consume extra processing time, and for that, the quality of the result might decline [4]. By removing these irrelevant and unimportant features, the classification performance can be improved [5]. Three well-known filter-based feature selection techniques, namely chi-square, feature importance, and relief, and their ensemble have been applied in this work. Our proposed approach works by minimizing the effect of a skewed distribution of classes in the training data to increase the value of true positive rate and classification performance.

The rest of the paper is organized as follows. After reviewing related work in Sect. 2, a brief overview of the dataset used in this study is discussed in Sect. 3. Then, the proposed methodology is given in Sect. 4, while in Sect. 5 provides the evaluation results through experiments and comparative analysis. Finally, the conclusion and the future scope are presented in Sect. 6.

2 Related Work

Dataset balancing concept is being used effectively for the imbalanced dataset. Kamei et al. evaluated the effect of over and undersampling on imbalance fault-prone module detection [6]. Ramezankhani et al. [7] applied SMOTE oversampling on imbalanced diabetes dataset to study the impact. Pelayo et al. [8] focus on the SMOTE oversampling to determine to improve the recognition of defect-prone modules. Huda et al. [9] used different oversampling technique to balance software defect prediction dataset. Zohu et al. [10] utilize the combined sampling technique which includes SMOTE oversampling and undersampling to balance the dataset. The feature selection has been performing effectively on software fault prediction since the beginning until now. Many approaches to feature selection have been used to bring efficiency to the software fault prediction sector. Agarwal et al. [11] propose a feature selection-based Linear Twin Support Vector Machine (LSTSVM) model to predict software modules that are defect prone. Xu et al. [12] in their paper investigate the impact of 32 feature selection techniques on the defect prediction performance. Shivaji et al. [13] have investigated multiple feature selection techniques that are generally applicable to classification-based bug prediction methods. With the use of five feature selection techniques, they discarded less important features. The findings and limitations of some previous works are summarized in Table 1.

3 Dataset Overview

The datasets used in the experiments herein are collected from the NASA MDP repository software engineering databases [25]. These datasets vary in the number of instances, their degree of complexity, and Imbalance Ratio (IR). In Fig. 1, a short overview of the two datasets used in this paper is shown. The imbalance ratio varies from 12.2% (highly imbalanced) to 1% (only slightly imbalanced). The diversity of the dataset is also considered in the number of instances, while CM1 has 344 instances, PC2 dataset contains 1585 instances.

Only the class attribute is categorical, the rest of the attributes contain numerical data. Both datasets contain 38 attributes, and all attributes are common.

4 Proposed Model for Fault Detection

Numerous fault detection methods have been studied over the years, and increasing the detection accuracy remains to be an important criterion for developers. However, when the dataset is imbalanced, and a single class makes up a significant part of the dataset, merely using accuracy as a metric is not sufficient. If data size differs greatly between major and minor classes, then detection rate of minor classes is heavily

Table 1 Comparative analysis of the state of the art

References	Findings	Limitations
[14]	Authors have applied five feature selection techniques on software defect prediction datasets to reduce the dimension	Data balancing was not discussed or performed
[15]	Authors have addressed different topics in which feature selection plays a crucial role	
[16]	Seven feature selection techniques were applied to select the most relevant features from software defect prediction dataset	
[12]	Two versions of NASA dataset were used in their study. One was noisy, and the other one was clean dataset. And 32 feature selection method was used	
[17]	Four filter feature ranking and fourteen filter subset selection methods were evaluated in this study	
[18]	Selection of attribute with log filtering was applied for feature selection, and naive Bayes classifier was used for defect prediction	No data balancing was applied, and only one classifier was considered to predict the software defects
[19]	Authors have used bat-based search algorithm for feature selection and resampling to balance the dataset	Only one classification algorithm was used for the defect prediction purpose
[10]	Their model adopts chi-square test for attribute selection, combination oversampling and undersampling to balance the dataset and J48 classifier for classification	Barely one feature selection method was used to select relevant features, and one classification algorithm was used for classification
[20]	Authors developed a hybrid sampling technique to balance dataset then compared their technique with SMOTE and virtual oversampling	The dimensionality of the dataset was not reduced
[21]	SMOTE oversampling was used to balance the dataset, and six classifiers were used for the classification	
[22]	Ensemble classifier was used for classification, and SMOTE oversampling was used to tackle imbalance distribution	
[23]	A fuzzy-based oversampling algorithm was used to handle the imbalanced data	
[9]	Authors built an ensemble model using different oversampling techniques, which consider the class imbalance problem	PROMISE software engineering dataset was used, and feature selection was not applied

(continued)

Table 1 (continued)

References	Findings	Limitations
[24]	A resampling technique with three types of ensemble learners was introduced to balance the dataset	

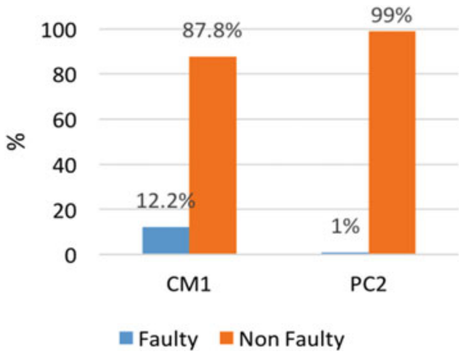


Fig. 1 Dataset overview [25]

affected; therefore, in this study, the skewness of the dataset is focused to remove and improve the recognition of the minority class or true positive rate.

To reduce the effect of imbalanced data in fault detection, the proposed model is implemented with original data according to the flowchart shown in Fig. 2. Techniques used such as the Z-score technique for the data cleaning step, SMOTE for data balancing, chi-square, relief, feature importance, and ensemble for feature selection step contributes to the elimination of outliers, minimization of the effect of asymmetry between classes and reduction of multi-dimensionality of the dataset, respectively. All the selected learning algorithms cover distinct types of methods in the training of imbalanced dataset.

All steps along with the techniques used have been briefly discussed below.

4.1 Data Cleaning

To reduce the possibility of noise generation, data cleaning has been applied in this model. It improves the quality of the training data by removing noise and correcting inconsistencies in the data.

First, the instances of the majority class that contain outliers have been identified, using the Z-score technique. The Z-scores [26] can quantify the unusualness of observation or in other words determine extreme points numerically. Instances that have a higher absolute value than the normalization threshold of 3 or -3 is recognized. This threshold value is commonly used for large datasets to detect and eliminate noisy

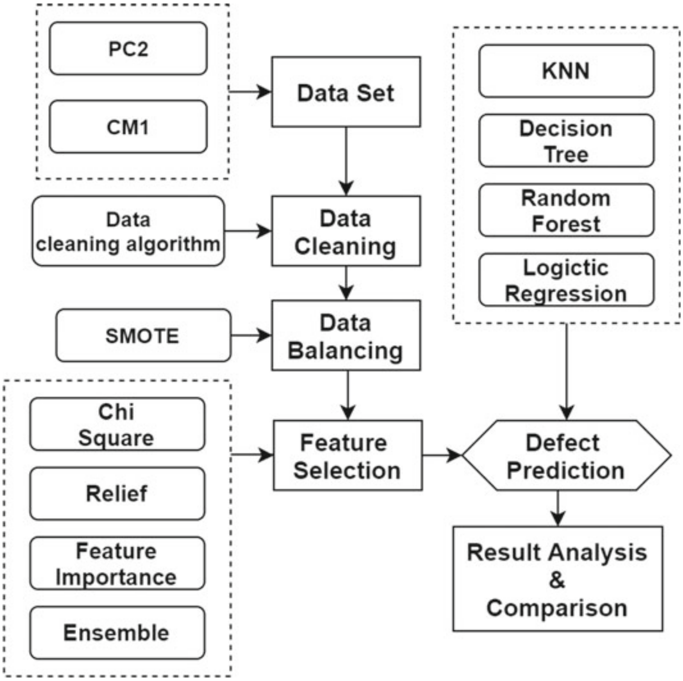


Fig. 2 Process mapping

data, where 99% of the values have a *Z*-score between 3 or −3; this means that they lie 3 standard deviations above and below the mean.

(a) Parameters for calculating *Z*-score

Z = Z-Score value

x = each value from the dataset

μ = mean or average

σ = standard deviation

n = total instances in the majority class

threshold = 3.

(b) Formulas for calculation *Z*-score

$$\mu = \frac{\sum_{i=1}^{n} x^{i}}{n} \tag{1}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} \left(x^{i} - \mu\right)^{2}}{n}} \tag{2}$$

$$Z = \frac{x^{i} - \mu}{\sigma} \tag{3}$$

To data cleaning, only the instances of the majority class have been considered, as the number of instances of the minority class is significantly lower. Any reduction from the minority class may lead to the generation of a high percentage of artificial data when oversampling will be performed. The pseudo-code of the proposed algorithm is illustrated in Algorithm 1.

```
1. Algorithm Begin
2. Separate the majority instances form the dataset
3. n← total instances in majority class
4. outlier_count←{0}
5. for i=1 to total_feature_in_dataset:
       average[i]←average of all x in feature i
       S.D[i]←standard deviation of each feature i
   End for
6. for i=1 to total_feature_in_dataset:
       for j=1 to n:
           Z←($x_j$-average[i])/S.D[i]
           If(abs(Z)>=threshold):
               outlier_count[j]+ +
       end for
   end for
7. for i=1 to n:
       if outlier_count[i]=0:
           remove $i^{th}$ instance form the dataset.
   End for
8. End of algorithm
```

In Algorithm 1, first, the instances of majority class are separated from the dataset, and the count value of the total number of majority instances is stored in variable *n*. An array is defined for outlier count, and it is initialized with the default value, zero. The index number of an array denotes the instance number. Therefore, outlier_count[*i*] indicates the total number of outliers in the *i*th instance.

Next, the average value and the standard deviation are calculated for each feature and stored in two different arrays named average[] and S.D[]. Here, the index number of the arrays denotes the feature number. Hence, average[*i*] means the average value of the *i*th feature and S.D[*i*] denoted the standard deviation value of the *i*th feature.

Then, for each feature, the *Z*-score value of every instance is calculated, and the absolute value of *Z*-score is compared with the threshold value. If the *Z*-score value is above or equal to the threshold value for jth instance, then the outlier_count[j] is increased by one. And finally, the instances, which have the outlier count greater than or equal to zero, are eliminated from the dataset.

4.2 Data Balancing

SMOTE oversampling generates synthetic instances in minority class to balance the dataset. This technique involves two key steps. In the first step, it searches k samples that are closest in distance to the minority class samples. The usual value of k is 5 [7]. In the final step, new synthetic instances are generated by the following procedure:

The difference between the minority class sample variable (x_i) and its nearest neighbor (x_j) is calculated. Then, the distance between x_j and x_i is multiplied by a random value between 0 and 1 finally added to variables value of minority sample.

$$\boldsymbol{x_{new}} = \boldsymbol{x_i} + \left(\boldsymbol{x_j} - \boldsymbol{x_i}\right) \times \boldsymbol{\delta} \tag{4}$$

Here, $\boldsymbol{x_{new}}$ is the generated synthetic instance, and $\boldsymbol{\delta}$ is the random value between 0 and 1.

4.3 Feature Selection Techniques

In the next step, feature selection is carried out. The applied feature selection algorithms used in the study are briefly discussed below.

4.3.1 Chi-Square

The chi-square [10] feature selection tests the independence of two events. The occurrence of a feature or attributes is considered as the first event, and the second is the occurrence of the class. It is a nonparametric test, where the hypothesis of no association between two or more groups is tested.

$$x^2 = \frac{(O_i - E_i)^2}{E_i} \tag{6}$$

In this formula, O_i represents the occurrence of attributes, and E_i represents the occurrence of the class. These two factors are involved in the calculation for a rank.

4.3.2 Feature Importance

Feature importance refers to techniques that assign a score to each input feature based on their usefulness to predict the target variable, and based on that score, more privilege is given toward the output variable of the features with higher scores than those with lower scores. It works by using the importance score to select features that are to be kept (highest scores) and those that are to be removed (lowest scores). It decreases the dimensionality of the dataset and thus simplifies the problem that is to be modeled. To compute the relative importance for all the attributes individually, an ensemble of decision trees is used by this technique [14].

4.3.3 Relief

The relief algorithm inspired by instance-based learning [27, 28] was formulated by Kira and Rendell [29, 30]. It works by finding the conditional dependencies between attributes of a dataset. During preprocessing, the relief algorithm is viewed as a feature subset selection algorithm that selects only those features that are statistically relevant to the target class. Relief requires linear time for the number of training instances and the number of supplied features irrespective of the target concept to be learned.

To help with the selection of relevant attributes, relief tends to find the sample size and a threshold of relevancy from the given training data. A total number of triplets is collected by relief. These are its near miss instance and near hit instance. To select near miss and near hit, Euclidean distance is used. Relief also calls a routine for updating weight vectors of features for every triplet and finds an average feature weight vector that is relevant to all the features to the target concept. The features that contain the average weight which is above the given threshold are selected by relief [5].

4.3.4 Ensemble

An ensemble of feature ranking techniques is an approach, where multiple feature ranking lists are obtained from corresponding feature ranking techniques [31], and all the obtained lists, one single ranking list is generated. This process takes place in two steps. First, a set of ranking lists is obtained with help of various feature selection algorithms. Second, a function is generated from a combination of functions which transform the ranking lists in step 1 into one individual list. The second step is more important as it includes the combining method.

This can be demonstrated formally. Consider a dataset D consisting of Y attributes with X records, the initial step would be to obtain a set of X ranking lists $\{F_1, F_2, \ldots F_n\}$, and the next step would be to work out a combination method T; let f_j^i indicate the rank of feature i from ranking list j, such that the set of rankings of feature i is given by $S_i = \{f_i^1, f_i^2, \ldots f_i^n\}$. Using the combination method T, the new score obtained by feature *I* is

$$f_i^{'} = T\left(f_i^1, f_i^2, \ldots f_i^n\right) \tag{6}$$

4.4 Classification Algorithms

For the classification process, four widely used classifiers have been considered for evaluating the efficiency of the feature selection methods. These are K-Nearest Neighbor (KNN) [32], logistic regression [33], decision tree [34], and random forest [35]. Table 2 gives a short description of the classification algorithms based on their characteristics and parameter settings.

5 Results and Analysis

In this study, the performance of data balancing and feature selection algorithms in fault detection procedures is examined. After data cleaning, majority class instances of CM1 are reduced to 266 from 302, whereas for PC2 dataset the majority instances become 1242 from 1569. The parameters are selected by default in all the implemented algorithms except for SMOTE and feature selection algorithms. In SMOTE, the percentage parameter for CM1 was determined to be 532%, and for PC2, it was 7662%. The percentage parameter says how many synthetic instances are created based on the number of the class with less instances. Moreover, the parameter numToSelect of the feature selection algorithms was set to 20. This resulted in 20 most significant features to be selected from each dataset.

All the experiments have been carried out by applying ten-fold cross-validation. This is to decrease the variability of the performance results due to the random generation of train and test sets. Weka 3.8.3 [36] was used on a desktop PC with 3.6 GHz Intel Core i7-4790 processor and 16 GB RAM to perform the experiments.

Table 2 Machine learning algorithms with their description and parameter settings

Classifiers	Description	Parameter settings
KNN	An instance-based classification algorithm	$K = 5$, NNSearch = LinearNNSearch
Decision tree	Tree-based classification technique	Confidence factor = 0.25
Random forest	Ensemble-based learning method for classification	nEstimators = 100
Logistic regression	Function-based classification technique	Kernel = PolyKernel; lambda = 0.01

Table 3 Performance results based on SMOTE after feature selection (20 features)

Dataset	Classifiers	Performance					
		Accuracy		TPR		TNR	
		Before balance (%)	After balance (%)	Before balance	After balance	Before balance	After balance
CM1	KNN	83.721	88.329	0.286	0.943	0.887	0.823
	Decision tree	84.593	85.687	0.262	0.879	0.927	0.835
	Random forest	86.046	91.713	0.024	0.925	0.977	0.910
	Logistic regression	86.628	79.473	0.214	0.785	0.957	0.805
PC2	KNN	97.918	98.591	0.063	0.998	0.989	0.974
	Decision tree	98.864	98.873	0.00	0.990	0.999	0.987
	Random forest	98.927	99.355	0.00	**0.**990	0.999	0.990
	Logistic regression	98.864	94.645	0.188	0.965	0.997	0.928

In Tables 3 and 4, the experimental results for the evaluated classification algorithms are tabulated in terms of accuracy, TPR, and TNR. These metrics are calculated according to Eqs. (7)–(9). To calculate these matrices, four important counts are collected from the confusion matrix: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Accuracy refers to the percentage of correctly classified defects.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \tag{7}$$

True Positive Rate (TPR) or sensitivity is the rate of actual positives that are correctly identified.

$$\text{TPR} = \frac{TP}{(TP + FN)} \tag{8}$$

True Negative Rate (TPR) or specificity is the rate of actual negatives that are correctly identified.

$$\text{TNR} = \frac{TN}{(FP + TN)} \tag{9}$$

In the classifiers, defects are considered as positive class, and non-defects are considered as negative class.

Table 4 Performance results based on feature selection (20 features) after applying SMOTE

Dataset	Classifiers	Performance					
		Accuracy		TPR		TNR	
		Before feature selection (%)	After feature selection (%)	Before feature selection	After feature selection	Before feature selection	After feature selection
CM1	KNN	89.266	88.323	0.947	0.932	0.838	0.835
	Decision tree	88.89	88.135	0.868	0.894	0.910	0.868
	Random forest	93.408	93.032	0.947	0.943	0.921	0.917
	Logistic regression	78.907	81.167	0.785	0.826	0.793	0.797
PC2	KNN	98.671	97.918	0.998	0.998	0.976	0.970
	Decision tree	98.752	98.832	0.990	0.986	0.986	0.991
	Random forest	99.396	99.519	0.997	0.997	0.991	0.994
	Logistic regression	95.330	95.491	0.969	0.977	0.938	0.933

Due to the limited scope of the paper, some results have been omitted from the presentation.

5.1 Comparison With no Balancing

As explained in Sect. 1, the initial datasets reflect a skewed class distribution. Every machine learning algorithm encounters a significant challenge when working with these higher number of defective classes and its high imbalanced instances. For the performance evaluation of our proposed approach, experiments have been carried out, with and without data balancing, among two preprocessed sets with feature importance feature selection.

Table 3 summarizes the performance of the SMOTE algorithm based on the four classifiers. It is indicated that the TPR of CM1 and PC2 dataset increased significantly. For both CM1 and PC2, the random forest algorithm is the most successful algorithm with a TPR increase of approx. 0.9 while decision tree and KNN being the second most efficient. This is followed by other algorithms applied to both the datasets. Similarly, it can be observed from Fig. 3 that after performing SMOTE, the TPR increase of approx. 0.99 can be obtained for PC2 dataset while applying a different feature selection algorithm. However, the accuracy of PC2 and CM1 shows a small change before and after SMOTE is applied. The variation in TNR is also very slight

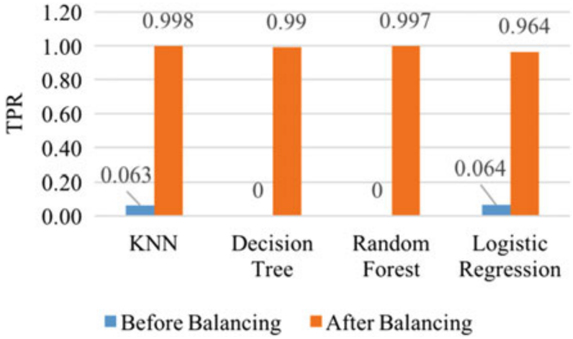


Fig. 3 Comparison of performance based on the selected attributes by ensemble feature selection of PC2 dataset before and after applying the SMOTE technique

and therefore negligible. This does not affect the goal of this paper as the focus was to predict faulty software correctly. Identifying all data points as non-faulty in the software fault detection problem is not helpful and instead it is important that we concentrate on identifying the positive class. The metric our knowledge tells us to maximize, TPR or the ability of a model to find all the relevant cases within a dataset.

5.2 Comparison with no Feature Selection

A comparison between before and after applying feature selection algorithms has been made to evaluate the performance of the proposed SDP. As a result of the selection of relevant features by the chosen attribute selection methods, the values of the metrics have increased slightly.

Table 4 presents the accuracy, True Positive Rate (TPR), and True Negative Rate (TNR) of the classifiers before and after the relief algorithm has been applied.

The results show that the performance of the classifiers is nearly similar before and after feature selection. This is also further proven in Fig. 4, where identical experiments have been carried out but with ensemble feature selection.

5.3 Comparison with Other Feature Selection Methods

To further evaluate our approach, a comparison between all the feature selection techniques are executed based on the CM1 and PC2 datasets and the random forest classification technique.

Although it is a common practice to apply the concept of ensemble learning to classification, other fields of machine learning, such as feature selection can also be improved by using it [37]. Machine learning methods alone can no longer deal with datasets efficiently as they increase in dimensionality and get more complex.

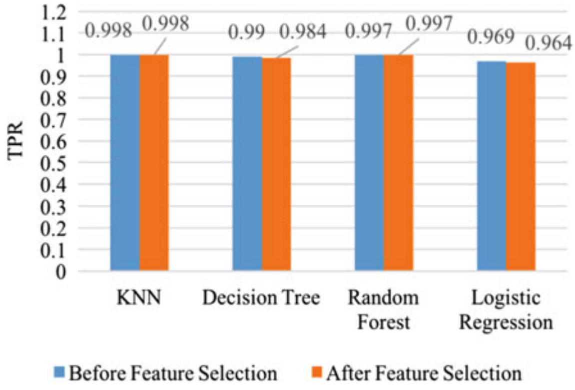


Fig. 4 Comparison of performance based on balanced PC2 dataset before and after applying ensemble feature selection

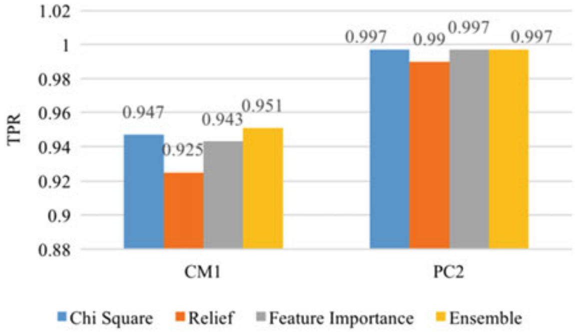


Fig. 5 Comparison performance of different feature selection techniques based on balanced datasets

As shown in Fig. 5, ensemble feature selection achieves the highest TPR on both CM1 and PC2. Therefore, in terms of performance and efficiency, it has superiority over other feature selection algorithms.

6 Conclusion and Future Scope

In automated software fault detection, balanced data which results in high accuracy, high TPR, and a high TNR is an important requirement. However, in real-life situations, most datasets are highly imbalanced and complex. The effect of low TPR and TNR can grow if not detected early in the software cycles. Hence, this issue must be addressed. In addition to balancing data, dimensionality reduction by choosing useful features can improve the efficiency of the classifiers and decrease training time.

This article shows the effect of oversampling and feature selection techniques on the results of four different machine learning algorithms when dealing with high-dimensional and imbalanced data. First, data cleaning is applied to minimize noise generation. Second, an oversampling technique, SMOTE is applied to both CM1 and PC2. This is followed by the application of four different feature selection techniques on each algorithm. The experimental results reveal that even though feature selection on its own has some effect albeit minimal but after combining oversampling and feature selection, the performance of the classifiers has improved significantly. Such results confirm the strength and usefulness of the proposed approach. Our approach outperforms other approaches that take only feature selection into account. Most importantly, because of initial data cleaning in the proposed method, can be balanced and apply any feature selection on noise-free datasets.

An apparent limitation of this study is that it examines only two datasets, but for better analysis of the performance in terms of consistency, the suggested method should be evaluated with other datasets. Moreover, only a small group of filter-based feature selection techniques and SMOTE oversampling method was considered. Filter-based feature selection methods ignore the dependencies between the features, and they also dismiss interaction with the classifiers.

This study considers only a small group of feature selection techniques and an oversampling method. In future work, investigating filter and wrapper-based feature selection techniques along with a combination of undersampling and oversampling methods might prove important. Also, instead of considering only a fixed number of features (in our case, 20), conducting experiments with a varying number of features could be beneficial.

References

1. Kaur R, Sharma ES (2018) Various techniques to detect and predict faults in software system: survey. Int J Futur Revolut Comput Sci Commun Eng (IJFRSCE) 4(2):330–336
2. Ali MM, Huda S, Abawajy J, Alyahya S, Al-Dossari H, Yearwood J (2017) A parallel framework for software defect detection and metric selection on cloud computing. Cluster Comput 20(3):2267–2281. https://doi.org/10.1007/s10586-017-0892-6
3. Ben-Gal I (2005) Outlier detection
4. Chandrashekar G, Sahin F (2014) A survey on feature selection methods. Comput Electr Eng 40(1):16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024
5. Vege SH (2012) Ensemble of feature selection techniques for high dimensional data
6. Kamei Y, Monden A, Matsumoto S, Kakimoto T, Matsumoto K (2007) The effects of over and under sampling on fault-prone module detection
7. Ramezankhani A, Pournik O, Shahrabi J, Azizi F, Hadaegh F, Khalili D (2016) The impact of oversampling with SMOTE on the performance of 3 classifiers in prediction of type 2 diabetes. Med Decis Mak 36(1):137–144. https://doi.org/10.1177/0272989X14560647
8. Pelayo L, Dick S (2007) Applying novel resampling strategies to software defect prediction. In: Annual conference of the North American fuzzy information processing society—NAFIPS, pp 69–72. doi: https://doi.org/10.1109/NAFIPS.2007.383813
9. Huda S et al. (2018) An ensemble oversampling model for class imbalance problem in software defect prediction. In: IEEE Access 6:24184–24195. doi: https://doi.org/10.1109/ACCESS.2018.2817572
10. Zhou L, Li R, Zhang S, Wang H (2018) Imbalanced data processing model for software defect prediction. Wirel Pers Commun 102(2):937–950. https://doi.org/10.1007/s11277-017-5117-z

11. Agarwal S, Tomar D (2014) A feature selection based model for software defect prediction. Int J Adv Sci Technol 65:39–58. https://doi.org/10.14257/ijast.2014.65.04
12. Xu Z, Liu J, Yang Z, An G, Jia X (2016) The impact of feature selection on defect prediction performance: an empirical comparison. In: Proceedings of ISSRE '96: 7th international symposium on software reliability engineering, ISSRE, pp 309–320. doi: https://doi.org/10.1109/ISSRE.2016.13
13. Shivaji S, James Whitehead E, Akella R, Kim S (2013) Reducing features to improve code change-based bug prediction. IEEE Trans Softw Eng 39(4):552–569. doi: https://doi.org/10.1109/TSE.2012.43
14. Tasnim Cynthia S, Rasul MG, Ripon S (2019) Effect of feature selection in software fault detection BT—multi-disciplinary trends in artificial intelligence, 2019, pp 52–63
15. Bolón-Canedo V, Sánchez-Maroño N, Alonso-Betanzos A (2016) Feature selection for high-dimensional data. Prog Artif Intell 5(2):65–75. https://doi.org/10.1007/s13748-015-0080-y
16. Jakhar AK, Rajnish K (2018) Software fault prediction with data mining techniques by using feature selection based models. Int J Electr Eng Inform 10(3):447–465. doi: https://doi.org/10.15676/ijeei.2018.10.3.3
17. Balogun AO, Basri S, Abdulkadir SJ, Hashim AS (2019) Performance analysis of feature selection methods in software defect prediction: a search method approach. Appl Sci 9(13). doi: https://doi.org/10.3390/app9132764
18. Sharmin S, SAL: an effective method for software defect prediction, pp 184–189
19. Ibrahim DR (2017) Software defect prediction using feature selection and random forest algorithm, pp 252–257. doi: https://doi.org/10.1109/ICTCS.2017.39
20. Zengin A (2016) HSDD : a hybrid sampling strategy for class imbalance in defect prediction data sets
21. Karatas G (2020) Increasing the performance of machine learning-based IDSs on an imbalanced and up-to-date dataset. IEEE Access 8:32150–32162. https://doi.org/10.1109/ACCESS.2020.2973219
22. Alsawalqah H, Faris H, Aljarah I, Alnemer L. Hybrid SMOTE-ensemble approach Adv Intell Syst Comput 1. doi: https://doi.org/10.1007/978-3-319-57141-6
23. Liu S, Zhang J, Wang Y, Xiang Y (2016) Fuzzy-based feature and instance recovery, pp 605–615. doi: https://doi.org/10.1007/978-3-662-49381-6
24. Abdou AS (2018) Early prediction of software defect using ensemble learning : a comparative early prediction of software defect using ensemble learning : a comparative study. doi: https://doi.org/10.5120/ijca2018917185
25. Gray D, Bowes D, Davey N, Sun Y, Christianson B (2012) Reflections on the NASA MDP data sets. IET Softw 6(6):549–558
26. Kannan KS, Manoj K, Arumugam S (2015) Labeling methods for identifying outliers. Int J Stat Syst
27. Aha DW, Kibler D, Albert MK (1991) Instance-based learning algorithms. Mach Learn 6(1):37–66
28. Callan JP, Fawcett T, Rissland EL (1991) CABOT: an adaptive approach to case-based search. IJCAI 1991(12):803–808
29. Kira K, Rendell LA (1992) A practical approach to feature selection. In: Sleeman D et al (eds) Morgan Kaufmann, San Francisco (CA), pp 249–256
30. Kira K, Rendell LA (1992) The feature selection problem: Traditional methods and a new algorithm. AAAI 2:129–134
31. Rahman A, Verma B (2013) Ensemble classifier generation using non-uniform layered clustering and Genetic Algorithm. Knowl-Based Syst 43:30–42. https://doi.org/10.1016/j.knosys.2013.01.002
32. Yihua Liao VR, Vemuri (2002) Use of K-nearest neighbor classifier for intrusion detection 21(5):439–448
33. Peng C-YJ, Lee KL, Ingersoll GM (2002) An introduction to logistic regression analysis and reporting. J Educ Res 96(1):3–14. https://doi.org/10.1080/00220670209598786

34. Quinlan JR (1986) Induction of decision trees. Mach Learn 1(1):81–106. https://doi.org/10.1007/BF00116251
35. Fawagreh K, Gaber MM, Elyan E (2014) Random forests: from early developments to recent advancements. Syst Sci Control Eng 2(1):602–609. https://doi.org/10.1080/21642583.2014.956265
36. Witten IH, Frank E (2002) Data mining: practical machine learning tools and techniques with java implementations. SIGMOD Rec 31(1):76–77. https://doi.org/10.1145/507338.507355
37. Pes B (2020) Ensemble feature selection for high-dimensional data: a stability analysis across multiple domains. Neural Comput Appl 32(10):5951–5973. https://doi.org/10.1007/s00521-019-04082-3

Indoor Navigation Assistant for Visually Impaired (INAVI)

U. B. Mahadevaswamy, D. Aashritha, Nikhil S. Joshi, K. N. Naina Gowda, and M. N. Syed Asif

Abstract Navigating in unknown places can be challenging for people who are deprived of the benefit of sight. The focus of advancements in navigation systems is extended onto helping the visually impaired understand the structure of their surroundings whilst they are traveling. The main idea behind the proposed method is to eliminate the dependency of the visually impaired on unreliable sources in an unfamiliar locality. Most of the public buildings these days, such as college and office buildings are equipped with their own Wi-Fi network which is used by the proposed indoor navigation assistant to direct the user while navigating inside the building. Using the Received Signal Strength Indicator (RSSI) values that are taken from each access point, trilateration is performed for localization, and the speech output guides the person by informing about the current location and thus making it possible for the visually impaired to move inside a building without any human assistance.

Keywords Wireless Fidelity (Wi-Fi) · Received Signal Strength Indicator (RSSI) · NodeMCU

U. B. Mahadevaswamy · D. Aashritha (✉) · N. S. Joshi · K. N. Naina Gowda · M. N. Syed Asif
Department of Electronics and Communication, JSS Science and Technology University, Mysuru, India
e-mail: aashutrisha@gmail.com

U. B. Mahadevaswamy
e-mail: mahadevaswamy@sjce.ac.in

N. S. Joshi
e-mail: nikhilsj98@gmail.com

K. N. Naina Gowda
e-mail: naiiinagowda@gmail.com

M. N. Syed Asif
e-mail: syedasifm.n@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_17

1 Introduction

Humans have been exploring and navigating around the world over centuries. Navigation systems, over the years, have evolved from maps and compasses to global positioning systems with detailed user interfaces, portraying the locations and routes with constantly improving accuracy. The realistic experience offered by current navigation systems adds on to the assistance provided by the same for finding places and directions to the same.

Even though these systems are very useful on road, they fall behind in providing the same facilities inside the four walls of a building. These systems are designed and modeled in such a way that they guide the user outdoors in unfamiliar places. However, this modeling has not been extended to indoors. The need for guidance inside unfamiliar buildings is tantamount to the requirement of directions while traveling around in a new city.

Assistance in moving around the corridors of a building is of most importance to the visually impaired people. The people deprived of the benefit of sight struggle to move around in an unfamiliar environment, especially without any guidance from others. An indoor navigation assistant can make their movement inside the building a lot comfortable by guiding them through the building's ways.

An indoor navigation system is a concept which is designed to assist the indoor movement to the user. The idea behind the indoor navigation assistant is to locate the user and guide through the ways of a building to reach the desired destination.

2 Related Work

Numerous research and implementations have been carried out on developing an efficient indoor positioning system that can be used worldwide. Several journal papers and research publications were considered for a better understanding of the field of this venture. The closest approaches have been studied in detail, and some of them are elaborated further.

In [1], Sunmin Lee, Jinah Kim, and Nammee Moon have proposed a smart watch which basically acts as a Wi-Fi-based indoor location recognition. The problem of position recognition due to the similar signal strength is solved by using both the Received Signal Strength Indication (RSSI) and Basic Service Set Identifier (BSSID).

The authors of [2] present an algorithm that a mobile device can utilize, as GPS-like reference nodes, either in range location-aware compatible mobile devices or pre-installed low-cost infrastructure-less location-aware beacon nodes.

The model proposed in [3, 4] uses neural networks to train the distance model. The authors propose an RSSI real-time correction method which is based on Bluetooth gateway.

In [5], the authors have built a Long Short-Term Memory (LSTM) recurrent neural network. This network makes regression between fingerprints and the locations in order to track the moving target.

Furthermore, indoor positioning has been an area of constant development with technologies ranging from IR to BLE included. In that view [6–9] discuss indoor localization using Bluetooth Low Energy (BLE) along with other additional protocols and techniques. The BLE beacons are deployed in different locations, and the different RSSI techniques are used for identifying the position of a user and mobile device.

A new technique of self-localization using infrared sensors is proposed [10]. The authors have come up with a configuration which consists of an IR LED array equipped with unique ID encoding capabilities which are based on a combination of different frequencies, and the repeated use of each ID encoding LED is done to address the issues of limited frequencies.

In recent times, Wi-Fi technology is preferred over other methods. The methods in [11, 12] provide some insight into the usage of Wi-Fi for indoor localization. Due to its wide deployment, Wi-Fi is expected to become a prominent tool for indoor positioning.

Along with this, Wi-Fi fingerprints can be further explored to identify human activities and locations. These conspicuous benefits of using Wi-Fi finger printing are well discussed in [13, 14]. Furthermore, [15, 16] provide details regarding the recent researches carried out in indoor positioning and the results of the same.

The existing indoor navigation methods use the maps and blueprints of the campus or buildings to guide the visitors. Some work has been done in developing various methods to determine the location of the visitors. An eclectic range of techniques have been discussed and proposed, explaining different methods of determining the position of the visitor by using the RSSI from the BLE beacons and Wi-Fi signals. Most of the works are limited to finding the position of the visitor.

Even though the ideas and experiments carried out so far have provided results to some level, they have either been conducted in a limited space area or have their own limitations with respect to the technology and approach. Thus, these previous works provide a good platform for the development of the existing methods.

3 Proposed Work

The system proposed intends to cover upon the area of indoor navigation. Further, specifically, under indoor navigation, the approach focuses on developing indoor navigation assistance for the visually impaired people.

In recent times, it can be seen that wireless technology has taken over the market completely. The indoor navigation assistant takes advantage of the available Wi-Fi network of a building to carry out the localization of the user, thus providing a novel facility by involving existing technology.

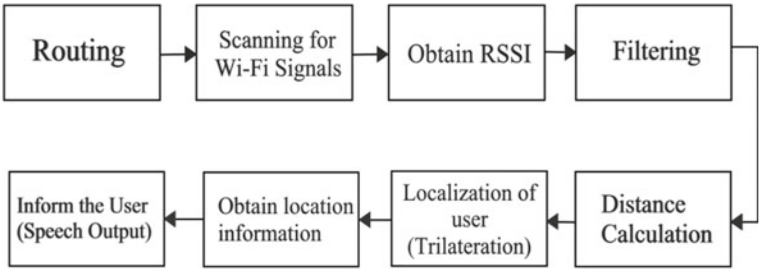

Routing
Scanning for Wi-Fi Signals
Obtain RSSI
Filtering
Distance Calculation
Localization of user (Trilateration)
Obtain location information
Inform the User (Speech Output)

Fig. 1 Proposed block diagram

Among the various algorithms used for localization, the most popular method is the Received Signal Strength Indicator (RSSI)-based algorithm. Received Signal Strength Indication (RSSI) is an indicator of the power level that the receiver sensors receive from access points. It is measured in terms of decibels from 0 (zero) to −120 (minus 120).

The implemented method utilizes Wi-Fi technology and trilateration techniques to determine the position of the user. Each user is provided with an ESP8266 device that is programmed to scan for the Wi-Fi signals and obtain their respective RSSI strength. Three Wi-Fi signals with the best RSSI are used to perform trilateration and obtain the distance of the user from the access points, thus obtaining the location of the user.

Once the location is obtained, the pre-fed database of the architecture of the building is searched, and the physical location details of the user are obtained. After this, the next important step is to inform the user about it. This is then achieved with the help of the speaker connected for audio output. In a new indoor environment, the visually impaired persons can survive independently without any human assistance. Figure 1 shows the block diagram which represents the working of the implemented system. The functionality of each block is as follows:

1. **Scanning for Wi-Fi signals**: The NodeMCU ESP8266 device is used to scan for the Wi-Fi signals from nearby access points.
2. **Obtain RSSI**: The Received Signal Strength (RSSI) of individual signals is obtained. The nearest three access points are chosen based on the strongest RSSI.
3. **Filtering**: The signals from the routers might undergo multi-path fading causing fluctuations in the RSSI. These fluctuations are controlled by the implementation of a Kalman filter.
4. **Distance calculation**: Once the RSSI is obtained, the distance of the device from the access points is calculated.
5. **Obtain (x, y) coordinates**: Coordinates of the three chosen access points are obtained from the pre-defined database.
6. **Localization of user**: Trilateration algorithm is used, and the device's location is determined.
7. **Obtain location information**: With the determined location, the information regarding that location is obtained from the pre-defined database.

8. **Inform the user**: The information fetched from the database is communicated to the user; thus, the user is aware of his\her surroundings using speech output.

The proposed method utilizes the Wi-Fi technology and trilateration techniques in order to determine the position of the user. Each user is provided with a NodeMCU ESP8266 device that is programmed to scan for the Wi-Fi signals. The device obtains the Received Signal Strength Indication (RSSI) from at least three Wi-Fi access points. RSSI is basically the measurement of power that is present in a received radio signal.

In case, if the device receives signals from more than three access points, then the top three signals with higher RSSI values are considered. To avoid the fluctuations in RSSI from affecting the calculations, a Kalman filter is used before the RSSIs which are considered for trilateration. The response time is affected by RSSI if the device fails to receive any Wi-Fi signals, thereby continuing its search for the signals in a loop. Using the RSSI, the distance of the device from the access point is determined using the formula:

$$\mathrm{RSSI(dbm)} = -10\,\mathrm{m}\ \log(d) + R \tag{1}$$

where

$R =$ signal strength of received signal in dBm.

$d =$ distance between the device and the access point.

$m =$ propagation constant or path loss (for free space: $m = 2$).

The NodeMCU ESP8266 device computes the distance "d" from the mentioned formula using three different signal strengths. To determine the position, trilateration algorithm is used. Trilateration is a technique used to determine the coordinates of a point using the distance of that point from three known locations. Using three distances, trilateration can pinpoint the exact location.

In Fig. 2, the signals coverage from the access points forms three circles that intersect each other at different points. When the device is in the intersection area, using the distances calculated before the location of the device can be determined by solving the system of linear equations given below. The following five steps are conducted to calculate (x, y) coordinates:

Step 1: The three respective equations for each of the three circles are as follows:

$$(a - x_1)^2 + (b - y_1)^2 = d_1^2 \tag{2}$$

$$(a - x_2)^2 + (b - y_2)^2 = d_2^2 \tag{3}$$

$$(a - x_3)^2 + (b - y_3)^2 = d_3^2 \tag{4}$$

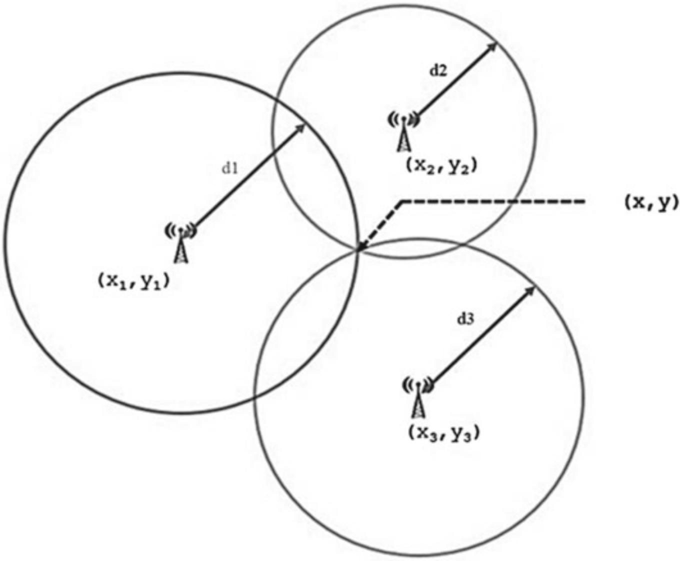


Fig. 2 Position coordinates detection

where

$a = x$ coordinate of the user.

$b = y$ coordinate of the user.

$x_1 = x$ coordinate of the first access point.

$x_2 = x$ coordinate of the second access point.

$x_3 = x$ coordinate of the third access point.

$y_1 = y$ coordinate of the first access point.

$y_2 = y$ coordinate of the second access point.

$y_3 = y$ coordinate of the third access point.

$d_1 =$ distance between the user and the first access point.

$d_2 =$ distance between the user and the second access point.

$d_3 =$ distance between the user and the third access point.

Step 2: Expanding out the squares in each of the above three equations

$$a^2 - 2x_1a + x_1^2 + b^2 - 2y_1b + y_1^2 = d_1^2 \tag{5}$$

$$a^2 - 2x_2a + x_2^2 + b^2 - 2y_2b + y_2^2 = d_2^2 \tag{6}$$

$$a^2 - 2x_3a + x_3^2 + b^2 - 2y_3b + y_3^2 = d_3^2 \tag{7}$$

Step 3: Subtracting Eq. 6 from Eq. 5:

$$(-2x_1 + 2x_2)a + (-2y_1 + 2y_2)b = d_1^2 - d_2^2 - x_1^2 + x_2^2 - y_1^2 + y_2^2 \tag{8}$$

Similarly, subtracting Eq. 7 from Eq. 6:

$$(-2x_2 + 2x_3)a + (-2y_2 + 2y_3)b = d_2^2 - d_3^2 - x_2^2 + x_3^2 - y_2^2 + y_3^2 \tag{9}$$

Step 4: Rewriting Eqs. 8 and 9 from step 3 using M, N, O, P, Q, R values:

$$Ma + Nb = O \tag{10}$$

$$Pa + Qb = R \tag{11}$$

Step 5: The solution of this system is:

$$a = \frac{OQ - RN}{QM - NP} \tag{12}$$

$$b = \frac{OP - MR}{NP - MQ} \tag{13}$$

The coordinates of the checkpoints including important laboratories, seminar halls, staff rooms, and staircases in the floor plans are determined by the implemented system and stored in the database by the campus admin. After the coordinates of the user are acquired, they are compared to that of the checkpoints in the database, thus obtaining the location of the user. The next important step is to convey this location information to the user.

The ESP8266 audio library comes in hand while generating speech output from the text with the support of ESP8266SAM library. Using a single transistor amplifier circuit with a speaker at the end, the speech output can be obtained. The transistor emitter is grounded, and the transistor base is driven by the ESP8266-I2SOUT (Rx) pin. The collector terminal and the USB 5 V supply are connected to negative and positive terminals of the speaker as shown in Fig. 3.

Even the 3 V from ESP8266 can provide the power but the volume is comparatively low. Since the ESP8266 pins cannot provide the necessary current, the amplifier is essential. Without that, there is a possibility of the device being damaged. Using an object of the audio output I2SNoDAC class, the speech output is produced from the text in the program. The location names are previously stored in the program which is converted into speech when the user is near that location.

Fig. 3 Connection between the ESP8266 module and the speaker

```
                            2N3904 (NPN)
                            +---------+
                            |         |     +-|
                            | E  B  C |    / S|
                            +-|--|--|-+    | P|
                              |  |  +------+ E|
                              |  |         | A|
ESP8266-GND ------------------+  |  +------+ K|
                                 |  |      | E|
ESP8266-I2SOUT (Rx) -----/\/\/\--+  |      \ R|
                                    |       +-|
USB 5V -----------------------------+
```

4 Result Analysis

The proposed system was implemented and was tested for its functionality inside the college building. Multiple trials were carried out to check the working of user localization. The testing was carried out by navigating from a source location to a destination location inside the building.

For testing, the system is designed to guide the user while navigating within the three floors of the college building. The architectural structures of each of these floors are depicted using Figs. 4, 5, and 6.

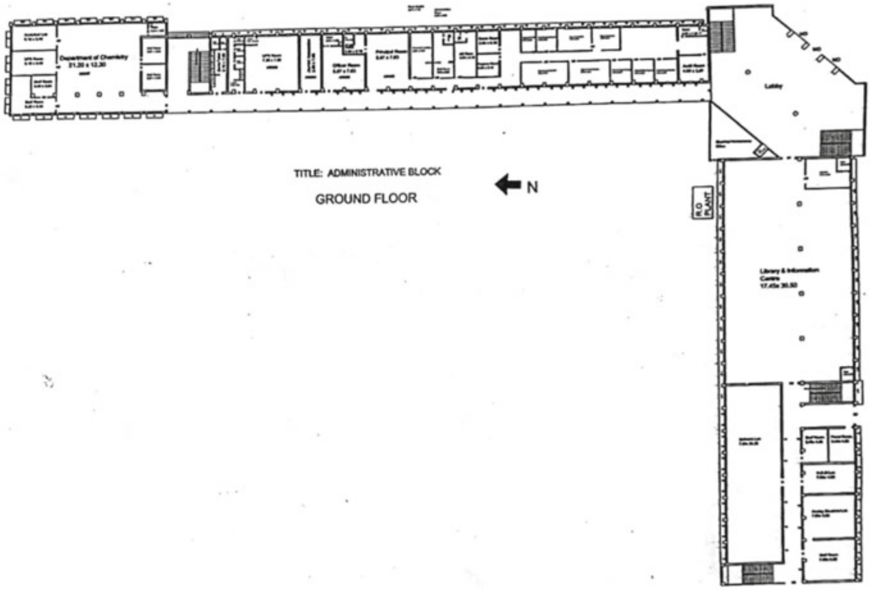


Fig. 4 Schematic of the ground floor

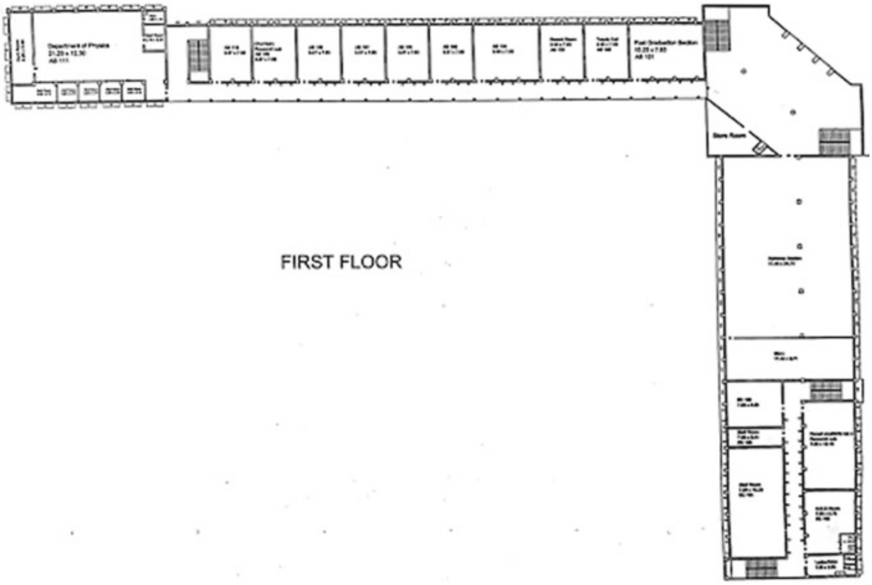


Fig. 5 Schematic of the first floor

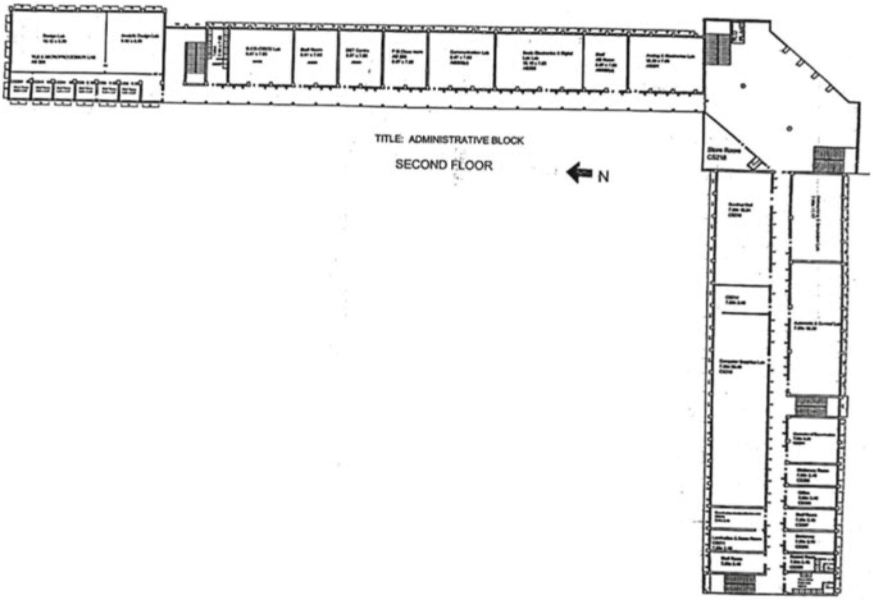


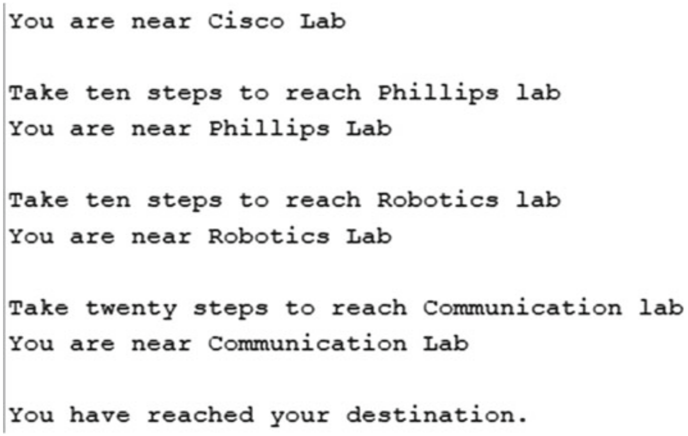
Fig. 6 Schematic of the second floor

With the inclusion of knowledge of staircases interconnecting all three floors, the system is equipped to work in the tortuous structure of the building. The system was tested for its accuracy in delivering correct location information, and the time taken to intimate the user was also noted over multiple trials.

Once the NodeMCU determines the user's location, it informs the user with a speech output using the ESP8266SAM library. The user becomes aware of his/her location once he/she reaches the checkpoints. Figure 7 shows a snippet of guiding directions that are conveyed to the user. However, in real time, the user is directed by similar speech commands.

The NodeMCU ESP8266 device functions quite effectively while dealing with this amount of data. The delay obtained while using the system is not very significant and is found to be under tolerable limits. It was observed that in real time, the device was successful in identifying the checkpoints inside the building (rooms or halls inside the building) as per the floor plans. Table 1 provides the time taken by the device to inform the user about his/her location over multiple trials.

Furthermore, the system's efficiency is analyzed by the accuracy with which the user's location is determined and conveyed to the user. At every stage of database



```
You are near Cisco Lab

Take ten steps to reach Phillips lab
You are near Phillips Lab

Take ten steps to reach Robotics lab
You are near Robotics Lab

Take twenty steps to reach Communication lab
You are near Communication Lab

You have reached your destination.
```

Fig. 7 Output commands from INAVI device

Table 1 Time taken to inform user

Trial no.	Database strength (No. of checkpoints)	Response time (seconds)
1	10	2.193
2	20	2.197
3	30	2.198
4	40	2.198
5	48	2.199
Average response time (seconds)		2.197

strength, 25 trials were done, and the trials, where the location was correctly identified, were considered a success. The trials wherein the location identification was either delayed or erroneous amounted to the error.

The error in measurement is calculated using equation

$$\%\text{Error} = \frac{|\text{Actual value observed} - \text{Expected value}|}{\text{Expected value}} \times 100 \tag{14}$$

$$\%\,\text{Accuracy} = 100 - \%\,\text{Error} \tag{15}$$

Using Eqs. 14 and 15, the accuracy at each database strength level was calculated, and in the final stage when all the checkpoints were included in the database, the system showed a location recognition accuracy of 96%. Even though the effect of varying database strength is negligible on the response time, an improvement in the accuracy can be seen. With the increase in the number of checkpoints, it was observed that an average localization accuracy of 1.5 m was achieved.

Wi-Fi signals are known to undergo fading, causing fluctuations in the RSSI. Therefore, to achieve accurate localization of the user not only the RSSI was used in the proposed method, but also the BSSIDs were utilized. The plot of the accuracy of the device, when tested over varying database strength, is shown in Fig. 8.

The implemented system was compared with that of the other systems implemented, and the findings are indicated in Table 2. The key difference when it comes to the system implementation and the complexity of the previously implemented systems and the drawbacks found in them are observed.

The system is found to be efficacious in dynamically tracking the user and identifying the user's position inside the building and guide him/her with further directions with voice output. This way it can keep the visually impaired user informed about

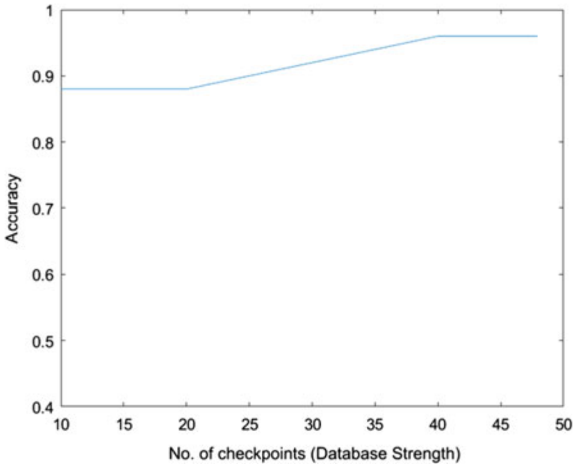


Fig. 8 Plot of accuracy with respect to the database strength

Table 2 Comparison with existing work

S. no.	Approach	Area of implementation	Key difference	System complexity	Result interpretation	Accuracy
1	RSSI random forest BSSID [1]	The system was implemented in open space (no walls in the surrounding)	Loss of signal strength and interference from other sources was not considered	Highly complex system due to the implementation of server and client approach for computations	Execution speed and accuracy results of the experiment are presented	Localization accuracy of 97.5%
2	RSSI DALIS [2]	The testing was conducted in an open space indoor environment with reference nodes deployed at three corners of area 4 × 4 m^2	The device was incapable of picking up low-level Wi-Fi signals	Simple setup but results in lagged consistency due to non uniform radiation patterns	The proposed method's architecture, scenario, localization, and accuracy for <3 m are compared with other methods	Localization accuracy is 97%
3	RSSI PSO-BPNN distance model BLE [3]	The implemented system did not consider the complex electromagnetic locations and the corners of the room	Noise affects the connection between the beacon and the smartphone	The involvement of BLE technology complicates the process of obtaining continuous positions with high precision	Comparison of positioning RMSE, MAE, and maximum error is done between the traditional and proposed method	Positioning error of 1.61 m No mention of accuracy percentage
4	Long short-term memory recurrent neuron network [5]	MATLAB and Spyder simulations with localization areas limited to 4 × $4m^2$	Only simulation results	Complex requires GPUs for fingerprints preparation	A comparison between LSTM, RNN, and BP is presented for training and testing time. Mean error and centralized time for different lengths in different trajectories are also presented	The mean error of different neural networks in different trajectories is plotted

(continued)

Table 2 (continued)

S. no.	Approach	Area of implementation	Key difference	System complexity	Result interpretation	Accuracy
5	INAVI	The system was implemented in a three story building	The system accesses the low-level Wi-Fi signals. Interference factors are also considered	Simpler, as computations are carried out only using NodeMCU ESP8266	Time is taken to inform the user, and the accuracy with respect to varying database strength is presented	Localization accuracy of 96% (an average accuracy of 1.5 m)

his/her position near the important landmarks of the building, thereby eliminating the dependency of the user on any other person.

5 Novelty

The proposed indoor navigation assistant has been developed with the main objective of helping the visually impaired inside the building. The proposed work not only locates the visitor but also intimates him/her with the information regarding his/her surroundings. Unlike the existing applications, the proposed idea does not demand the user to have the map of the building since it uses the pre-defined database with important locations of the building. The method presented here employs the trilateration technique instead of triangulation and thus does not require the orientation of the device and access points. The proposed idea uses minimal hardware that includes ESP8266 and a speaker prominently. Most of the existing works limit themselves to indoor localization whereas the idea proposed here is an extension of the former, and the visitor is informed about the important locations such as rooms, cabins, and halls as he/she wanders inside the building. The proposed system can be developed with the deployment of undemanding technologies whilst providing efficacious results and serving the visually impaired.

6 Conclusion and Future Scope

The proposed system is designed and implemented such that it tracks the user dynamically and updates the user about his/her current position. The system can inform the user about his location once he reaches near the designated checkpoints (rooms or halls inside the building). This will make it possible for the user to be independent while navigating in an unfamiliar environment. This leads to eradicating the need of human assistance to the visually impaired person.

However, the system has the limitation that the device has to receive at least three signals since three coordinates are a requirement for trilateration. Therefore, the system might be incompetent in coverage of fewer areas. This can be remedied by relocating the Wi-Fi routers to avoid areas without reception. The system performance can be enhanced by the additional feature of speech interaction, where the user can enter the destination location using speech input. A more dynamic approach in tracking and assisting the user can be accomplished with the inclusion of guidance to prevent obstacles.

The system implemented is such that the speech output is available in the English language as a default. But if developed further, then the system can be customized to provide the speech output to the user in the local languages as preferred by the user.

Moreover, the response time of the system can be improved by using a processor with high processing speed so that the delay in localization can be minimized. The further improvised systems can be implemented using Bluetooth low-energy devices, which will further enable a reduction in power consumption when implemented on a larger scale.

References

1. Lee S, Kim J, Moon N (2019) Random Forest and WiFi fingerprint-based indoor location recognition system using smart watch. Human-centric Comput Inf Sci 9(6)
2. Awad F, Al-Sadi A, Al-Quran F, Alsmady A (2018) Distributed and adaptive location identification system for mobile devices. EURASIP J Adv Signal Process 61
3. Li G, Geng E, Ye Z, Xu Y, Lin J, Pang Y (2018) Indoor Positioning Algorithm based on the improved RSSI distance model. Sens J 18(9):2820
4. Robesaat J, Zhang P, Abdelaal M, Theel O (2017) An improved BLE indoor localization with Kalman-based fusion: an experimental study. Sens J 17(5):951
5. Xu B, Zhu X, Zhu H (2019) An efficient indoor localization method based on the long short-term memory recurrent neuron network. IEEE Access 7:123912–123921. IEEE
6. Mekki K, Bajic E, Meyer F (2019) Indoor positioning system for IoT device based on BLE technology and MQTT protocol. In: IEEE 5th world forum on internet of things (WF-IoT), pp 787–792. IEEE, Limerick, Ireland
7. Heyn R, Kuhn M, Schulten H, Dumphart G, Zwyssig J, Trsch F, Wittneben A (2019) User tracking for access control with bluetooth low energy. In: IEEE 89th vehicular technology conference (VTC2019-Spring), pp 1–7. IEEE, Kuala Lumpur, Malaysia
8. Qureshi UM, Umair Z, Hancke GP (2019) Indoor localization using wireless fidelity (WiFi) and bluetooth low energy (BLE) signals. In: IEEE 28th international symposium on industrial electronics (ISIE), pp 2232–2237. IEEE, Vancouver, Canada
9. Sawaby AM, Noureldin HM, Mohamed MS, Omar MO, Shaaban NS, Ahmed NN, El Hadidy SM, Hussein RS, Hassan AH, Mostafa H (2019) A smart indoor navigation system over BLE. In: IEEE 8th international conference on modern circuits and systems technologies (MOCAST), pp 1–4. IEEE, Thessaloniki, Greece
10. Wang J, Takahashi Y (2018) Indoor mobile robot self-localization based on a low-cost light system with a novel emitter arrangement. ROBOMECH J 5(17)
11. Yang C, Shao HR (2015) WiFi-based indoor positioning. IEEE Commun Mag 53(3):150–157
12. Ebner F, Fetzer T, Deinzer F, Grzegorzek M (2019) On Wi-Fi model optimizations for smartphone-based indoor localization. Int J Geo-Inf 6(8)

13. Wang F, Feng J, Zhaoi Y, Zhang X, Zhang S, Han J (2019) Joint activity recognition and indoor localization with WiFi fingerprints. IEEE Access 7
14. Molina B, Olivares E, Palau CE, Esteve M (2018) A multimodal fingerprint-based indoor positioning system for airports. IEEE Access
15. Liu Q, Qiu J, Chen Y (2016) Research and development of indoor positioning China communications. 2016(2z):67–79
16. Yazti DZ, Laoudias C, Georgiou K, Chatzimilioudis G (2017) Internet-based indoor navigation services. IEEE Internet Comput 21(4). IEEE Computer Society

On the Evaluation of Effectiveness of eLearning and Blended Learning

Sarka Hubackova

Abstract Currently, the development of information and communication technology is influenced by versatile information and increases the potential to analyze the information through digital technologies. Most of the learners utilize ICT as a tool to get enhanced vision in their carrier. The use of multimedia in teaching and learning is examined to explore research. The effectiveness of teaching is dealt with interactive teaching and the use of eLearning. In order to obtain the relationship between the students and teachers, ICT will be a better platform. This research work presents an analysis of issues in educational process and its effectiveness and provides possible solutions for eLearning to increase the effectiveness in face-to-face teaching learning process.

Keywords Effectiveness · Assessment · eLearning · Teaching · Learning · Foreign languages · Teaching methodologies

1 Introduction

In the present situation, everyone uses smart gadgets like mobile phones, tabs for accessing various applications through Internet. All these applications are data oriented and include various data processing features. The technology development brings multimedia into handheld device and influences the society through its communication modules and reduces the communication gap through its fast processing characteristics. This provides various new opportunities and opens up various ways in education systems. Effective and efficient education system are a complex process, and it requires various factors which are practically difficult to measure. Various research works are evolved to evaluate the eLearning and teaching practices, and among them, Kirkpatrick methodology of evaluation is much familiar.

It is a simple and common model, although it was created as early as in 1959. Later, it began to be used for evaluation of eLearning. Among contemporary authors

---

S. Hubackova (✉)
Department of Applied Linguistics, Faculty of Informatics and Management, University of Hradec Kralove, Hradec Kralove, Czech Republic
e-mail: sarka.hubackova@uhk.cz

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_18

dealing with the measurement and evaluation of effectiveness, the example of Khan or Syverson should be shown. Original view on the aspects of eLearning (CAPEODL) was created by Khan [1, 2].

In education, different types of effects can be seen such as training effectiveness, individual effectiveness, pedagogical effectiveness, education system effectiveness, education process effectiveness and school effectiveness.

2 Methods

Initially, a method of the literature review of available sources exploring the issue of modern teaching methods was used. Research work utilizes databases such as science direct, Scopus, Web of science and Springer. The research issues are identified by evaluating the research works. A questionnaire is framed to obtain the relationship between the students and ICT. Also to obtain the teaching supported by ICT in student's view. The research process is started during 2018–2019, and all the questions are provided with multiple choices.

3 Effectiveness of eLearning

In the following part of our paper, the effectiveness and its relation to eLearning have conversed, respectively. The basic assumption of the course is a kind of technical, and any student who might take part in an eLearning teaching process can receive their appropriate e-material and be excellent [3]. This presumption stands complex of pedagogical assumptions.

Unlike face-to-face teaching, eLearning is more turned to an individual. The academic result of it is that a tutor should have a certain and if possible, a clear idea about the language knowledge of the students attending the seminar group [4].

Nowadays, eLearning is widely adopted, and smartphones supports the teaching and learning process. Smartphones are easily available, and multimedia teaching through smartphones gains more attraction in the recent days [5].

A tutor can add to a student's interests but cannot rely on the attractiveness of the new method. Also, an attractive and newly treated eLearning material are very important here; it supports student's attention and concentration, and in this way, it becomes a good assumption of eLearning effectiveness.

An important role in effectiveness may be played also by a suitable time extent in which eLearning is used. If this use should be effective, then it cannot be too long. So, if the contents of eLearning material and new information should be more complicated, then it is better to divide the instruction into two or three shorter intervals. Their effectiveness is higher than the efficiency of a one long time stage.

eLearning offers a good possibility of checking the quality of a student's cooperation. The grade of effectiveness depends here not only on the intensity of student's work, but can also be influenced by a teacher's attitude. It is necessary to mention in this connection a certain tutor's self-criticism and his willingness to revise the contents of the given task, its extent, the sequence of its items that the student must observe. eLearning makes it possible to inform the tutor about all of it. Such information is usually also a picture of the immediate effectiveness of the tutor's activity. Any change in the procedure or improvisation is here difficult. It is therefore important to think of a possibility of change during the tutor's home preparation. The tutor has to count on the possibility to change the method and use blended learning.

The necessity of making short inputs with closed content directed if possible, to a simple problem has its close connection with a time limitation of eLearning. In accordance with its complexity, the tutor decides if the relevant foreign language is used and assigns the tasks or explains the instruction in the students' mother tongue. Sometimes, it is advantageous to explain the instruction in the foreign language but to repeat its succinct contents in the mother tongue.

The effectiveness of the eLearning method is higher in the case when the task with its all details is assigned in a way that excludes further students' questions on how to do this or that. It is valid here in principle: The more distinct the assignment of a task is, the higher the measure of the effectiveness of the used method will be. The specification of tasks has also its psychological importance: A student takes for superfluous any help of another person, and he embarks upon his task with interest and force.

It is not suitable to use eLearning partially for presenting of information only on the one hand and for assigning of tasks on the other hand. Any one-sidedness here has a calamitous psychological influence, and it reduces the measure of effectiveness.

Both the variants have just mentioned can be accompanied by a very short and simple text. It does not necessarily have the nature of an examination. Sometimes, the tutor wants to check that the student understands an instruction or a task well. In another case, the assigned task applied to the information or instruction explained earlier. The test is not only a depiction of the short-time effectiveness of tutor's attitude, but it can show a student some faults in his work. This short test opens the possibility to check the cooperation of a seminar group and its members with a tutor. The awareness that he has such a possibility usually forces students to higher effort. But in our experience, the test assigned in this way is not the suitable source material for obtaining a credit. A face-to-face way for a much more suitable form is taken in this case.

4 Effectiveness of the Process of Teaching

The concept of effectiveness of teaching as a didactical category is not unambiguously delimitated. Contemporary literature discussing this pedagogical field proves this fact quite clearly. The terms effectiveness, efficiency as synonyms were used. In this sense, the effectiveness is seen as a measurable trace of a teaching process, a trace having a certain duration in the cultural consciousness of an individual who completed a teaching process. [4, 5].

The sub-terms used in our explanation are tried to explain with some occasional examples from different fields of teaching. The similar trace that has been in mind when mention a teaching process can be left in an individual's cultural awareness also by some other process. One of them is, for example, reading: a contact with individuals from other national fields or cultures, the experience of life or contact with educated members of the same nation often belonging to different generations. A special place is then taking here by different forms of self-study connected both with direct school teaching and with further lifelong education. The student's does not have school environment on the mind, but the educational occasions that can take their places in a gallery, a museum, on archaeological workplaces, in a workshop, in a concert hall or a laboratory are considered.

It is not possible to reflect on the full content and extent of a teaching process in this connection. There are, namely—almost at the same time—also two other processes: An individual, who accepts the content of a teaching process, makes with its perception concerning usually knowledge almost immediately a selection of both its content and its extent. But the facts that an individual keeps unwittingly or consciously in his consciousness are touched also by a quite different process, by a process of forgetting. Of course, a teaching process does not concern new information only, but also skills and teaching methods. And the process of forgetting treats the named entities in a quite different way than the pieces of information. Those plain facts are mentioned for two reasons: On one hand, wanted to point out that the effectiveness of a teaching process does not concern the mediation of pieces of information. On the other hand, wanted to emphasize that a face-to-face teaching must be taken into account and the more permanent effectiveness in which revision and strengthening of learned facts, their connection with other facts, etc. The duration of effectiveness which has been already mentioned in our definition represents a very complicated problem. To simplify it, immediate, short-term and long-term efficiency are distinguished. The share of one of them reveals most markedly in the situations, where it concerns the trace of face-to-face teaching. In the short-term and long-term effectiveness, the share of the effectiveness of teaching is always very high.

Also, some other factors come into effect: organization of learning and teaching of the relevant subject, its position in the educational plan of the relevant school purposefulness of the sequence of its contents, logical relations of separate basic elements of learning materials, used methods and books or teaching aids. Even the connection of effectiveness with the aim of the relevant subject in a certain span of time might shows itself as a very tight one. A teacher himself sometimes makes a self-examination about the study material and also in need of having a decision over the material. The important factors to be considered are with usefulness and applicability of the acquired facts.

The teaching in a foreign language lesson devoted for acquiring new vocabulary might have a very high immediate effectiveness. The teacher and the method can guarantee also it is extending to an acceptable level of short-time effectiveness and the focus on the new words from different points of view. The corresponding tutor chooses and explains the most important new words in connection with their frequency, meanings and grammar or orthography. And a certain kind of selection that have been mentioned. However, one can guarantee a prolongation of the immediate effectiveness into a short time one also by means of the textbook utilized. School textbooks usually remember the taught new words consistently and deliberately in a form of different exercises; they occasionally repeat at least the basic words of the new vocabulary. The teacher depends on the textbook in this case and does not pay any specific attention to newly educated vocabulary.

The basic relation mentioned here is to figure out the valid time. The intermediate effectiveness is a significant condition of the short-time efficiency, and this one is a condition of the long-time effectiveness.

The false students' opinion of the usefulness of the material should study usually is a big obstacle of the effectiveness of teaching. The tutor should judge the material that should be taught in the lesson from this point of view in his home preparation and prepare a suitable motivation. The purpose of the material and its place in the system of information and its indispensable place in the cultural consciousness of an individual.

The relation of methods and effectiveness leads to another speculation. One example is that a foreign language teaching students' mutual communication and conversation with the teacher play an important role. However, organic parts of a lesson can represent also some other activities, as, for instance, explanations concerning orthography or grammar rules that do not proceed as a communication nevertheless they can contribute to lessons' effectiveness evaluated in its relation to its scientific aim. Even such a modern method as eLearning does not necessarily contribute to higher effectiveness if its excessive use limits the process of authentic students' communication.

Motivation, often beyond the advancements and older students at universities always lead to a higher level of effectiveness. An interesting text attractive not only by its content but also by the tasks connected with its analysis can play a certain motivation in foreign language teaching. Even perfectly silent reading might be a very effective process. Teachers know well the situations when one can feel in the very quiet atmosphere of class that the students work with interest and very hard.

There is also a short of possibility of effectiveness' measuring in our definition. The effectiveness is usually measured by means of a test. A teacher does not measure his effectiveness in most cases but interested only in the extent or content of the taught material that the student masters. The harmful influence of frequent testing in this connection is to be mentioned. Students feel very soon what kind of information their teacher will check because he takes it for important. And this fact results in students' attitude to such information and their quite different attitude to the rest of the taught material. This student's attitude reflects the effect that even a very conscientious and scrupulous teacher's work leaves behind itself a very vague trace.

5 Findings

Contemporary pedagogy takes advantage of the most modern methodologies. A standard connectivity between eLearning and teaching is termed as blended learning [6, 9]. Through this blended learning, the satisfaction level of students with the teaching process could be furnished. Students responses for the survey is depicted as a graph below, and the feedback will help to find out the students rank and course complexity. In the survey, 78 students are allowed to participate, and out of 78, 72 students submitted the survey. The question framed was more suitable for face-to-face learning or learning supported by an online course. The students are asked about which method they most admire and most effective. Students considered blended learning to be the most effective method which is described in Fig. 1.

The importance of a teacher for effective teaching is also asked, and the role of the teacher is also considered by the students to be very significant which is explained in Fig. 2.

Figure 3—During the self-study, the following eLearning materials are used most often:

Dictaphone	1%
TV	4%
DVD/video	4%
Textbook	15%
Smartphone	18%
Internet	27%
Multimedia courses	31%

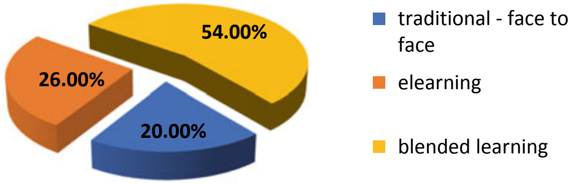


Fig. 1 Method of teaching

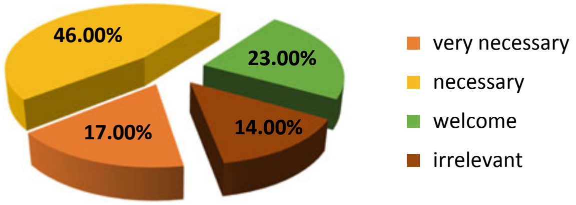
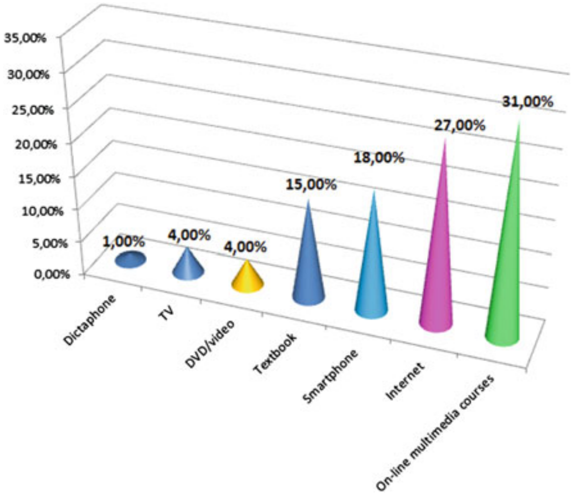


Fig. 2 Role of a teacher

Fig. 3 Most often used eLearning materials



6 Conclusion

The success of eLearning depends on many factors. The quality of the course as a whole, on the virtual environment in which the education takes place, on the readiness of students to work in a virtual learning environment, their ability to orient in the environment and in the use of all the tools that information and communication technologies offers, the personality of the tutor and the ability to run the course well and responsibly. And a very important role in the effective use of ICT in the education process is played by the student's attitude to individual work in the virtual learning space.

The eLearning represents potential possibilities for measuring both the method and the eLearning materials as Khan points out. But the contemporary pedagogic research heads towards another direction. The scholars do not talk about the effectiveness of a teaching process, but about the effectiveness of schools.

Acknowledgements This article was supported by the project titled SPEV (2104/2020) at the Faculty of Informatics and Management of the University of Hradec Kralove, Czech Republic.

References

1. Khan BH, Granato LA (2017) Program evaluation in eLearning. https://asianvu.com/digitallibrary/elearning/elearning_program_evaluation_by_khan_and_Granato.pdf. Last accessed 23 Sept 2017
2. Syverson MA, Slatin J (2017) Evaluating learning in virtual environments.https://www.learningrecord.org/caeti.html. Last accessed 18 Sept 2017

3. Poulova P (2003) Využití eLearningu ve vysokoškolské výuce z pohledu studentů a vyučujících. In: E-learning—Sborník příspěvků ze semináře a soutěže eLearning 2003, Gaudeamus Hradec Králové
4. Průcha J (2002) Moderní pedagogika. Praha
5. Skalková J (2007) Obecná didaktika. Praha
6. Frydrychova Klímova B, Poulova P (2011) Tutor as an important eLearning support. Proc Comput Sci 3:1485–1489. https://www.sciencedirect.com/science/article/pii/S1877050911000378. Last accessed 22 Feb 2011
7. Frydrychova Klimova B (2009) Blended learning. In: Research, reflections and innovations in integrating ICT in education, Lisboa
8. Hoffmann L (1987) Kommunikationsmittel Fachsprache: Eine Einfuehrung. 3. Aufl. Akademie Verlag, Berlin
9. Pikhart M (2014) New horizons of intercultural communication: applied linguistics approach. Proc Soc Behav Sci 152:954–957
10. Hubackova S (2010) Foreign language teaching with WebCT support. Proc Soc Behav Sci 3:112–115

Design and Performance Analysis of GaAs-Based P-i-N Photovoltaic Using AlGaAs as Window Layer

Rocky Chakma, S. S. Mahtab, M. J. Alam, and Rupa Akter

Abstract In this study, the simulation is based on the P-i-N solar Photovoltaic. P-i-N-based reference solar cell founded on GaAs for enhanced performance was utilized with AlGaAs as the window layer. Different layers have been optimized according to the best performance. Impact of using Anti Reflecting coating has also been studied. It has been found that with Anti Reflecting coating, P-i-N solar cells efficiency has increased significantly. A horizontal band section has been placed in the intrinsic substance of the P-i-N model and because of this, high η (efficiency) is build up with high short circuit current (Isc); high open-circuit voltage (Voc) as extra short circuit current density (Jsc) comes from the intrinsic section. 40.9133% is the highest η in the research.

Keywords P-i-N · GaP · Photovoltaic · Inorganic solar cell · GaAs

1 Introduction

Producing electrical power from light power is a one-step alteration by PV arrangement whose justification is dependent on quantum theory. Photons observed from light, whose power relies on frequency and shade of the light. Energy from photons is enough to electrify electrons and in the advanced power levels, they are freer to

---

R. Chakma
Department of EEE, USTC, Chittagong, Bangladesh
e-mail: rocky.cht@gmail.com

S. S. Mahtab
Department EEE, Feni University, Feni, Bangladesh
e-mail: mahtabshahzad@gmail.com

M. J. Alam (✉)
Department of EEE, Feni University, Feni, Bangladesh
e-mail: alameee1993@gmail.com

R. Akter
Department of EEE, Mymensingh Engineering College, Mymensingh, Bangladesh
e-mail: rupa.mec.eee.bd@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_19

be in motion [1]. In a p-i-n solar unit, an intrinsic section is inserted between p- and n-type coating of a GaAs p-n solar unit. For this case, Jsc is higher than the common p-n cell because of the supplementary carrier donation of the intrinsic section. As a result, efficiency increases. In quantum well solar units, the difference among the only bandgap and solar band's event power is provided by supplementary power like photons from diverse power can be concentrated well. So, it has an essential likeness. Here, mainly the concern is the production of various light-produced quasi-Fermi stages [2]. Among the two entrances, the transfer of the carriers is significantly unlike. In quantum well solar unit application, transport is achieved when the carriers at each sectional energy level escape due to light absorption. The runaway moment ought to be faster than recombination moment so that the accumulation of η can be obtained in the greatest number. Well and truly the quantum-well infrared photodetectors illustrate the expediency of the scape process [3], which have peak assemblage from intra-sub band process. Sectional band approaches have additional superiority in that consecutive sectional energy levels can have dissimilar energies, therefore authorizing a good quantity of productive band gaps and high performance [4].

2 Structure Analysis

The individual solar cell has a complex construction with various coatings over and above the p- and n-coating. The function of these coatings is to decrease face and reverse shell recombination along with superficial mirroring. A rough copy of a P-i-N indicated cell is shown in Fig. 1 that is screening every coating such as windows, anti-reflective coating, p^+ , P-i-N, and n^+) to get great η.

The V-I characteristic of the only p- and n-coating is presented. In this paper, the V-I characteristic of a solar cell configuration is derived from the continuity equation for both electrons and holes, with the correct boundary conditions for this construction. Inspired by [5] and the V-I quality of n^+-n-p junction is resultant.

2.1 Proposed Structure

In Table 1, various resources are used in suggestion cells.

For the opposition of reflective coating, R(E) is very much condensed also the reflection damage may become short around 2% [6]. In this research, the model has been used as a reflectivity equal to zero for every wavelength. The I-coating which is in among p- with n-coating is preferably un-doped. Having a grounding doping, N_i is greatly lesser than the injecting of p- and n-coating. Deficiency section expands in P-i-N connection addicted to the mild doped section. Within the p-i-n connection, depleted thicknesses of both p- and n-coating are understood as little [7]. Throughout a front outside field, greatly doped p^+-coating above the p-coating condenses the exterior recombination rate. P-layer's electrons get together a probable

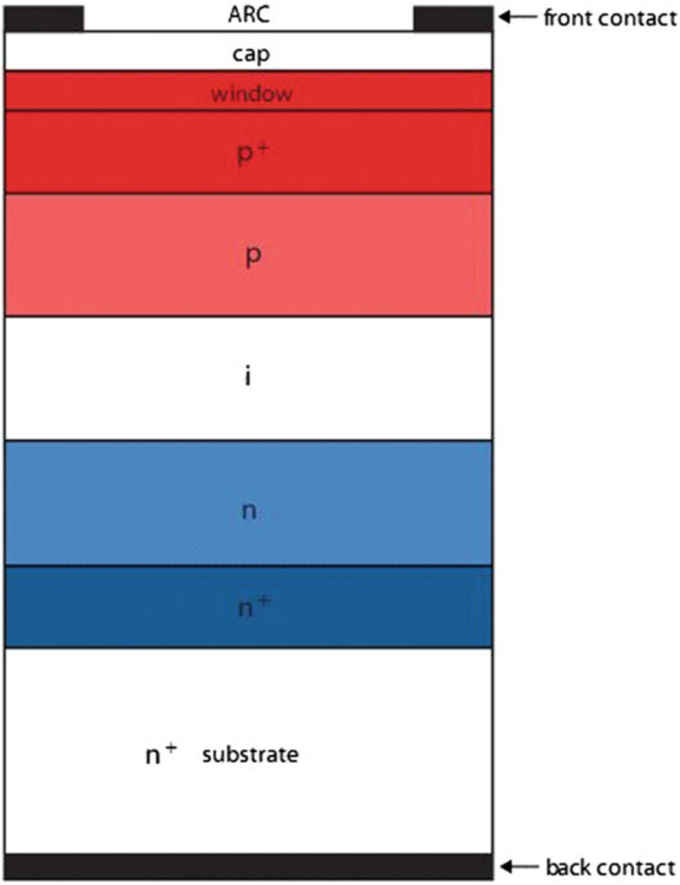
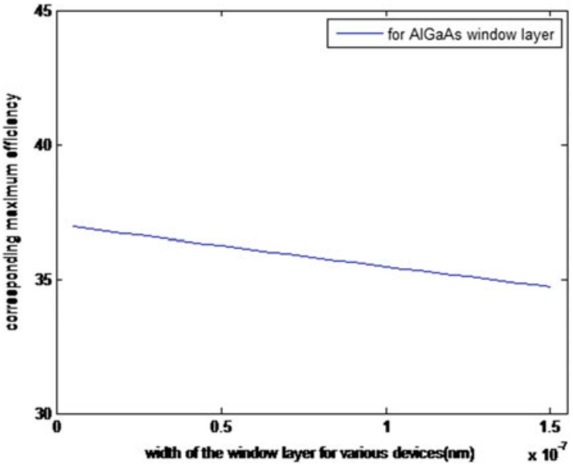


Fig. 1 P-i-N reference cell's construction

Table 1 Materials used in a suggestion cell

Coatings	Materials
Window	AlGaAs
p^+	GaAs
P	GaAs
I	GaAs
N	GaAs
n^+	GaAs

Fig. 2 Various efficiency for various thickness of $Al_xGa_{1-x}As$ ($x = 0.804$) window layer



barrier reliant on top of the proportion of doping intentness both in p$^+$-and p-coating while going as of p-coating hooked on the p$^+$ coating. It is known as front exterior field [8]. As the P$^+$ coating is located beside the front exterior where there is greatest photon flux, the carriers produced within p$^+$-coating supply much to the photocurrent [5]. In the boundary of un-depleted p-coating at z = 0, whole electron current (Fig. 2) is the addition of the donations. It is from the P$^+$ coating, p-coating along with depleted section among the P$^+$, and p-coating.

$$j_{n,\text{total}}(0) = j_{n,p}(0) + j_{n,p} + (0) j_{n,\text{dept}}(0)$$

Inserting a window coating, above the p$^+$-coating; which is prepared by advanced band gap substance, the valuable exterior recombination rate is more condensed. By placing supplementary greatly doped n$^+$ coating below the n-coating provides exterior field. It reduces the valuable rear exterior recombination rate similarly as the front exterior.

2.2 AlGaAs Seeing that Window Coating

Surplus carrier's recombination takes place not only inside the size of a semiconductor but also inside the exterior of it. By cutting short the periodicity at the exterior, it performs as boundary among the semiconductor and an additional substance. So, the surface recombination rate is different (mainly superior); which can be cut by passivating (or window coating) with the intention that it can avoid minority carrier as of going to the exterior, rather than the size of the semiconductor. Exterior recombination rate is robustly reliant on its roughness, pollution, ambient gases. In other material-based window layer such as GaAs, exterior recombination rate is excessive (of the order of cm/s). It can be cut capable of 10-cm/s by deposing a thin coating of AIGaAs [9].

3 Computational Methodology

3.1 P-i-N Solar Cell

3.1.1 State's Band Construction with Valuable Density

Al_xGa_{1-x} As and GaAs come together having the same lattice constants in zinc blende arrangement (variance is just 0.12%). Which indicates As that can be developed on GaAs with no damage can now stay away from the configuration of dislocations that might raise the recombination [10].

$$E_{\text{GaAs}}(T) = \left[1.519 - \frac{5.405 \times 10^{-4} T^2}{T + 204}\right] \text{eV}$$

For Al_xGa_{1-x} As smallest conduction band's location the Γ, X, and L bands are reliant on x, and As is a straight bandgap semiconductor for the condition of $x < 0.45$. It is shown away in bellow [11]

$$E_{\text{Al}_x}\text{Ga}_{1-x}\text{As} = E_{\text{GaAs}} + 1.247x$$

Al_xGa_{1-x} is for an indirect bandgap semiconductor for the condition of $x \geq 0.45$ with a bandgap written as

$$E_{\text{Al}_x\text{Ga}_{1-x}\text{As}} = 1.911 + 0.005x + 0.245x^2$$

Gallium arsenide's permittivity having temperature reliance is specified as follows:

$$\in = \left[12.79\left(1 + 1.0 \times 10^{-4} T\right)\right] \in_0$$

This equation for the extension of the depletion region in the moderately injected region which is utilized from [12] and manipulated as an estimation for w_a

$$w_a = \sqrt{\frac{2 \in k_B T}{q^2 N_a}} \arctan\left(\frac{q V_{pp} +}{k_B T} \sqrt{\frac{N_a^+}{2 N_a}}\right)$$

A fixed voltage of the n^+-n joint has the form [13]. On or after [12] the depletion region's length is in use hooked on the lightly doped n-section and utilized at the same time as an approximation for,

$$w_b = \sqrt{\frac{2\text{€}k_B T}{q^2 N_d}} \arctan\left(\frac{qV_{nn}+}{k_B T}\sqrt{\frac{N_d^+}{2N_d}}\right)$$

4 Numerical Results and Discussion

4.1 The Window Layer's Consequence of Thickness

For a variety of thickness of window coating, the highest η and Jsc have been estimated with the constant thickness of other coatings and further factors (Table 2).

From the data table, it is clear that increasing the width of the window layer causes to decrease efficiency because it increases recombination. As $\text{Al}_x\text{Ga}_{1-x}\text{As}$ (where $x = 0.804$) has an indirect bandgap, non-radiative recombination (i.e., Auger recombination) dominant here where radiative recombination is concealed because of the requirement of photons in this method. The usefulness of neither a pure initial substance nor a fresh fabrication procedure can improve this auger recombination. Only its lower thickness can cut the auger recombination rate in the interior of the cell. So, a lower thickness device gives better performance (Fig. 3).

This graph represents how efficiency is decreasing with the increase of windows layer thickness. At 5 nm thickness of windows layer, the peak efficiency can be calculated which is approximately 36.9481 (Fig. 4).

From the graph, the highest density can be obtained at 5 nm windows coating which is nearly 450 A/m^2 whereas the cut off voltage 0.9 v is estimated (Fig. 5).

This graph elucidates the decrease of quantum efficiency with the increase of wavelength at different windows layers. From the graph, it is more clear that 5 nm windows layer is much more effective. Although the other thickness shows nearly the same values but can be ignored. At 400 nm wavelength, the peak quantum efficiency

Table 2 Changeable thickness of window coating and simulation result

Device no	Window layer's thickness in nm	AlGaAs seeing that a window coating		
		Volt	A/m^2	Max efficiency
1	5	1.0054	445.1377	36.9481
2	10	1.0053	444.2197	36.8683
3	50	1.0046	436.9647	36.2371
4	100	1.0038	428.1129	35.4670
5	150	1.0029	419.4953	34.7172

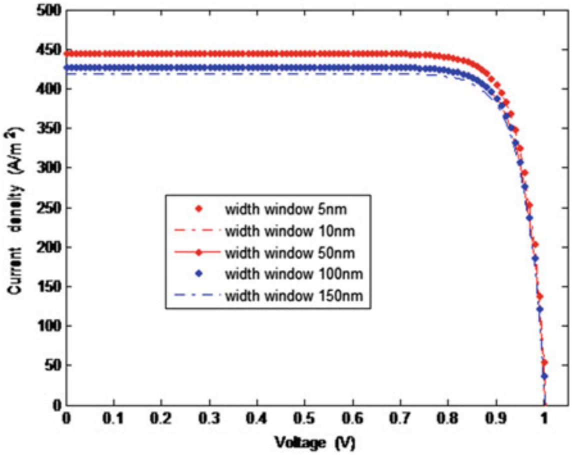


Fig. 3 V-I density arc for diverse width of the AlGaAs window coating

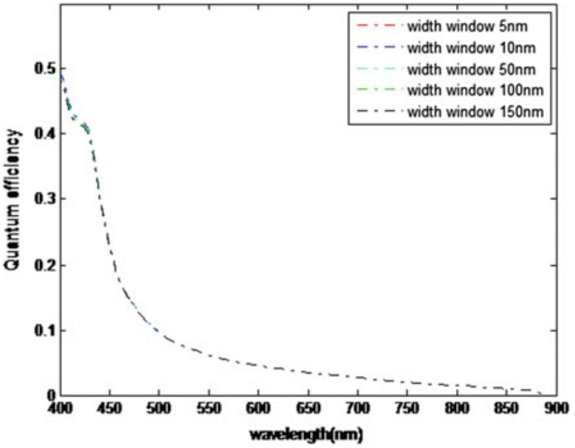


Fig. 4 λ versus QE in favor of diverse thickness of AlGaAs window coating

can be found. No light absorbed after the 900 nm wavelength because of the bellow bandgap of the active layer.

4.2 P^+ Coating's Consequences of Thickness

For a variety of thickness of p^+ coating, the highest η and Jsc have been estimated with the constant thickness of other coatings and further factors. To see the difference, AlGaAs window coating devices information is contained in Table 3.

A dark injected p^+-layer is found below the window layer, which additionally diminishes the front-surface recombination. The data table confirms that η and thickness of the p^+ layer are in addition linked. Thickness and η are proportional to each

Fig. 5 V-I density arc of the p$^+$ coating in favor of a range of thickness (via AlGaAs while a window coating)

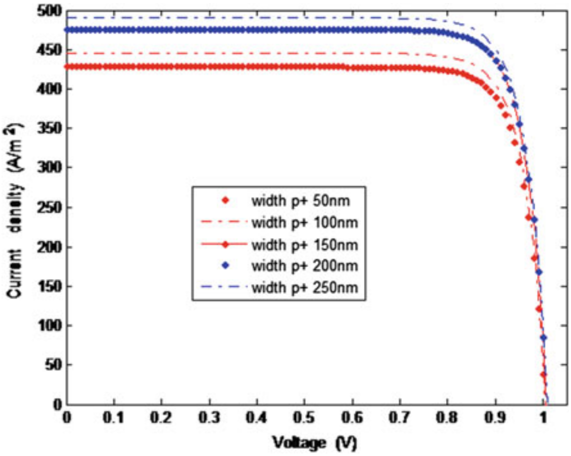


Table 3 AlGaAs window coating devices information

Device no	p$^+$ layer's thickness in nm	AlGaAs seeing that a window coating		
		Volt	A/m^2	Max efficiency
1	50	1.0038	428.4171	35.4935
2	100	1.0054	445.1377	36.9481
3	150	1.0068	461.0628	38.3385
4	200	1.0080	476.1407	39.6652
5	250	1.0092	490.3254	40.9133

other as a consequence of the absorption of more photon can be done in this coating. The data table also confirms that AlGaAs window coating has an identical mechanism presentation (Fig. 6).

The data from the graph directly represents that efficiency is also associated with the thickness of the p$^+$ layer which is also reflected in the data table. Enhancing the width of the p$^+$ layer also enhances the efficiency of the PV solar cell. This is happened because of the absorption of more photons by this layer. Considering the optimizing thickness, 250 nm p$^+$ layer has been calculated for the value of 40.9133% (Fig. 7).

This graph better designates how quantum efficiency enhances with the enhancing of the thickness of the p$^+$ layer considering optimized wavelength. The photocurrent is increased with the increase of quantum efficiency; As photocurrent is precisely associated with the quantum efficiency.

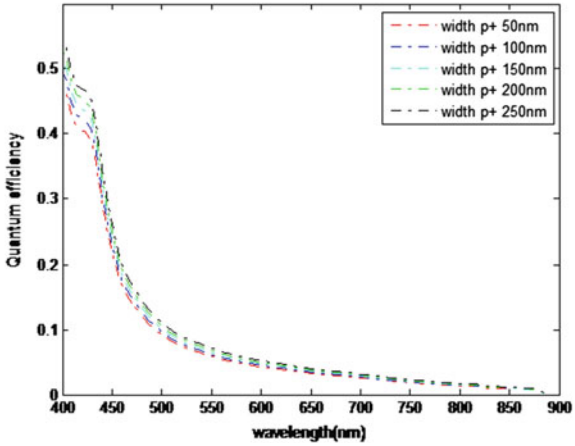


Fig. 6 λ versus QE in favor of diverse thickness of p+ coating (via AlGaAs while a window coating)

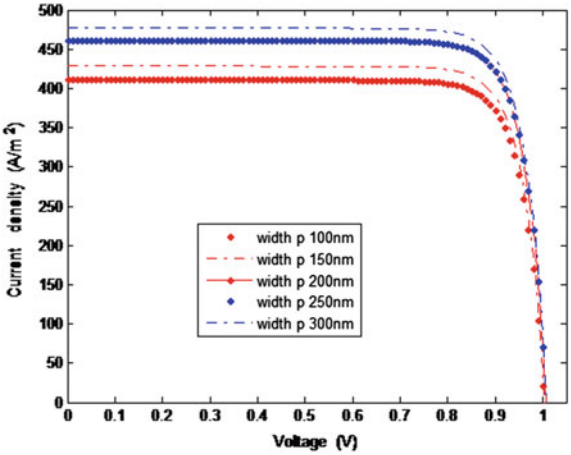


Fig. 7 V-I density arc of the p-coating in favor of a range of thickness (via AlGaAs while a window coating)

4.3 P-Coating's Consequences of Thickness

For a variety of thickness of p-coating, the highest η and Jsc have been estimated with the constant thickness of other coatings and further factors. To see the difference, AlGaAs window coating devices knowledge is carried in Table 4.

Various curves are obtained from a simulation that is given below (Fig. 8).

It can be seen that the current density increases when the thickness of p-coating increases. Considering the pv cell thickness, 300 nm P-coating has been optimized for 476.6671 A/m^2 current density (Fig. 9).

As the quantum efficiency directly associated with the thickness of P-coating, the quantum efficiency enhances with the enhancing of P-coating at an optimized

Table 4 For a changeable thickness of p-coating and simulation results

Device no	p layer's thickness in nm	AlGaAs seeing that a window coating		
		Volt	A/m^2	Max efficiency
1	100	1.0022	410.9044	33.9701
2	150	1.0038	428.3638	35.4890
3	200	1.0054	445.1377	36.9481
4	250	1.0068	461.2356	38.3536
5	300	1.0081	476.6671	39.7113

Fig. 8 λ versus QE in favor of diverse thickness of p-coating (via AlGaAs while a window coating)

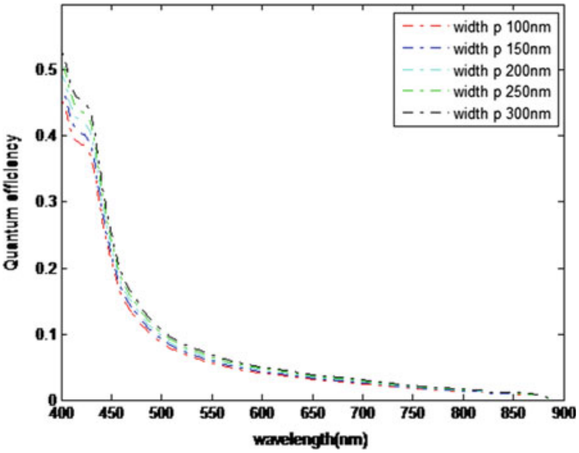
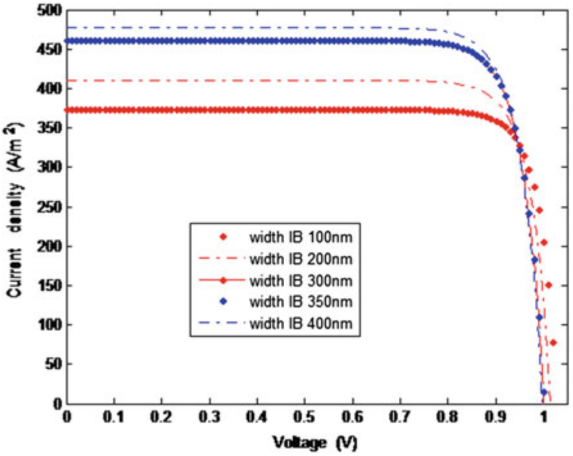


Fig. 9 V-I density arc of the intrinsic coating in favor of a range of thickness (via AlGaAs while a window coating)



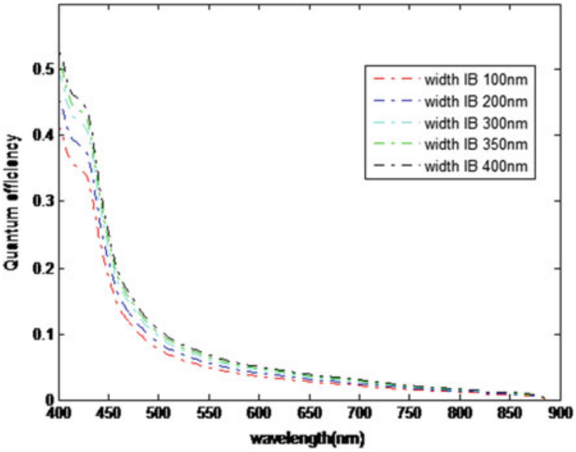


Fig. 10 λ versus QE in favor of diverse thickness of p-coating (via AlGaAs while a window coating)

wavelength which is reflected in the graph. It also can be noticed from the graph that the peak quantum efficiency can be obtained at 300 nm P-coating layer.

4.4 Intrinsic Coating's Consequences of Thickness

For a variety of thickness of the intrinsic coating, the highest η and Jsc have been estimated with the constant thickness of other coatings and further factors. The device information of the AlGaAs window coating is contained in Table 5.

Various curves are obtained from the simulation that is given below.

As an intrinsic layer increase the photon absorption, the short circuit current is increased with the increase of intrinsic layer and this is well demonstrated in the graph. 400 nm intrinsic layer has been optimized for the short circuit current of 477.2973 A/m^2.

Figure 11 confirms that QE is proportional to the thickness of the intrinsic coating. As a result, Isc too increases which is given in Fig. 10. The thickness is proportional to the number of photons soaked up, and power is also relative to the quantity of

Table 5 AlGaAs window coating

Device no	Intrinsic layer's thickness in nm	AlGaAs seeing that a window coating		
		Volt	A/m^2	Max efficiency
1	100	1.0281	373.4749	32.3951
2	200	1.0150	410.5894	34.6451
3	300	1.0054	445.1377	36.9481
4	350	1.0014	461.5055	38.0810
5	400	0.9979	477.2973	39.1655

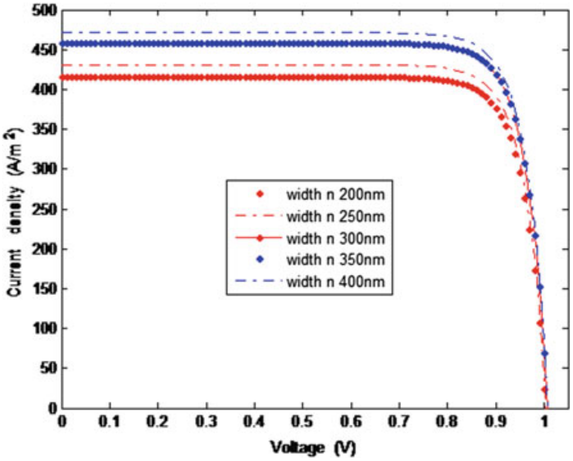


Fig. 11 V-I density arc of the n coating in favor of a range of thickness (via AlGaAs while a window coating)

photon soaked up so η is related to the thickness of the intrinsic coating. But if the thickness is sufficiently increased, power saturates.

4.5 N-Coating's Consequences of Thickness

For a variety of thickness of n-coating, the highest η and Jsc have been estimated with the constant thickness of other coatings and further factors. To see the difference, AlGaAs window coating devices knowledge is carried in Table 6 (Fig. 12).

From the graph, it can be noticed that the maximum short circuit current (Jsc) can be obtained at 400 nm depositing n layer whereas the minimum Jsc is at 200 nm.

The data table and arc confirm that thickness is proportional to η and Isc, because it soaks further photon.

Table 6 Changeable thickness of n-coating and simulation results

Device no.	Intrinsic coating's thickness in nm	AlGaAs seeing that a window coating		
		Volt	A/m^2	Max efficiency
1	200	1.0024	415.9974	34.4099
2	250	1.0039	430.9711	35.7150
3	300	1.0054	445.1377	36.9481
4	350	1.0067	458.5358	38.1171
5	400	1.0079	471.2188	39.2340

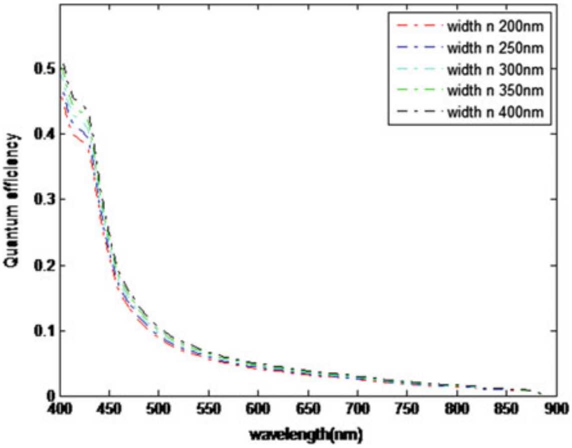


Fig. 12 λ versus QE in favor of diverse thickness of n-coating (via AlGaAs while a window coating)

4.6 n^+ Coating's Consequences of Thickness

To decrease the back surface recombination, a dark injected n^+-layer is accommodated beneath the n layer. For a variety of thickness of n^+ coating, the highest η and Jsc have been estimated with the constant thickness of other coatings and further factors. To see the difference, AlGaAs window coating devices knowledge is carried in Table 7.

The arc from Fig. 13 confirms that the highest Isc can be originated when the thickness of the n^+ coating is 150 nm, and the minimum can be found when the thickness of the n^+ coating is 25 nm. So, η increases proportionally with the increase of the depth of the n^+ coating. Figure 14 also confirms that maximum quantum efficiency can be originated when the thickness of the n^+ coating is 150 nm at an optimized wavelength.

Table 7 Changeable thickness of n^+ coating and simulation result

Device No.	Intrinsic coating's thickness in nm	AlGaAs seeing that a window coating		
		Volt	A/m^2	Max efficiency
1	25	1.0000	426.2351	35.2831
2	50	1.0022	433.4167	35.9173
3	75	1.0039	439.5884	36.4606
4	100	1.0054	445.1377	36.9481
5	150	1.0076	455.0994	37.8234

Fig. 13 V-I density arc of the n$^+$ coating in favor of a range of thickness (via AlGaAs while a window coating)

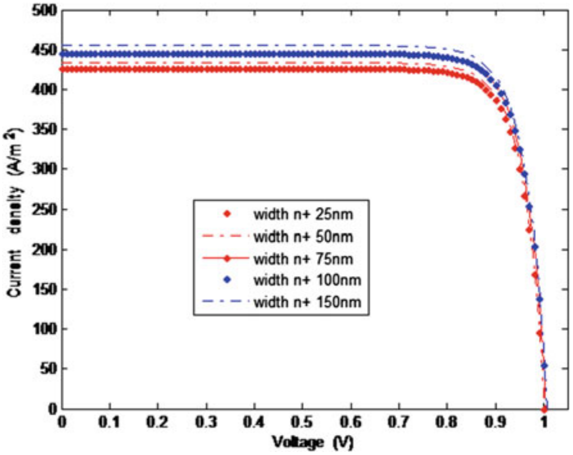
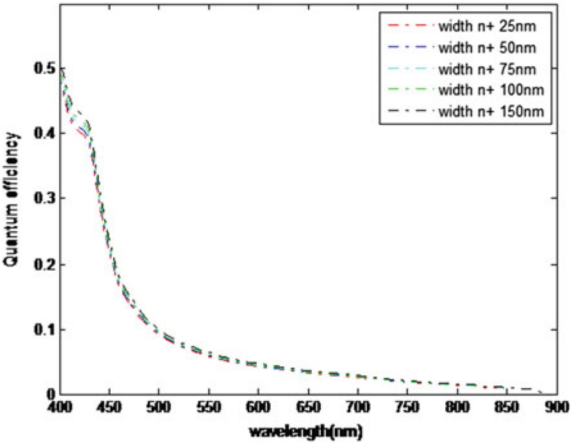


Fig. 14 λ versus QE in favor of diverse thickness of n$^+$ coating (via AlGaAs while a window coating)



5 Conclusion

In our research, the AlGaAs formed P-i-N indicated cell designed for leading presentation have been studied. To understand it, the P-i-N indicated cell's theoretical learning, as well as presentation, is vastly vital. Design of p-i-n indication cell is optimized by a simulation representation for great η from which preparation and presentation are put together. The indication cell's representation describes the consequence of a window coating. It also describes densely injected and coating for obtaining a small efficacious exterior alliance velocity with an anti-reflective coating to minimize the mirroring dissipation. Via this thickness, great QE is obtained. For the representation of P-i-N cell with the fundamental substance that is accommodated in a flat band area, a representation of great η is built up and high Isc and Voc is acquired because further Jsc comes from the intrinsic area. For AlGaAs substances, simulation is done with various factors particularly the thickness of diverse coatings. For AlGaAs, the greatest η is 40.913% for this research.

Conflict of Interest In this paper, all the authors have worked and contributed equally. The serial of author list included is basically according to seniority not by contribution. They all contributed equally and will be designated as the first author of this paper.

References

1. International Energy Outlook (2009) Energy Information Administration Office of Integrated Analysis and Forecasting U.S. Department of Energy, Retrieved online on 2009-06-13. http://www.jointsolarpanel.nl/fileadmin/jointsolarpanel/user/documents/seminar2004/stanleyzc04.pdf
2. Twidell J, Weir T (2006) Renewable energy resources. Taylor and Francis, Milton Park
3. Nelson J (2007) The physics of solar cells. Imperial College Press, London
4. Greenpeace and European Photovoltaic Industry Association (2008) Solar Generation V-2008, Solar electricity for one billion people and two million jobs by 2020
5. Dai XM, Tang YH (1996) A simple general analytical solution for the quantum efficiency of front-surface-field solar cells. Solar Eenergy Mater Solar Cel ls 43:363
6. Tobin SP, Vernon SM, Bajgar C, Geoffroy LM, Keavney CJ, Sanfacon MM, Haven VE (1988) Device processing and analysis of high efficiency GaAs cells. Solar Cel ls 24:103
7. Luque A, Martí A (2001) A metallic intermediate band high efficiency solar cell. Prog Photovolt: Res Appl 9:73
8. Martí A, Stanley CR, Luque A (2006) Intermediate band solar cells (IBSC) using nanotechnology, chapter 17 in Nanostructured Materials for Solar Energy Conversion. Elsevier B. V., Amsterdam
9. Martí A, Cuadra L, Luque A (2002) Quasi-drift diffusion model for the quantum dot intermediate band solar cell. IEEE Trans Electron Devices 49:1632
10. Streetman BG, Banerjee SK (2006) Solid state electronic devices. Pearson, London
11. Adachi S (1993) Properties of aluminium gal lium arsenide. INSPEC, 1993
12. Hauser JR, Littlejohn MA (1968) Approximations for accumulation and inversion space-charge layers in semiconductors. Solid-State Eelectron 11:667
13. Demoulin PD, Lundstrom MS, Schartz RJ (1987) Back-surface field design for n + p GaAs cells. Solar Cel ls 20:229

Impact Analysis of Anti-Reflection Coating on P-i-N Solar Device

S. S. Mahtab, Rupa Akter, Tauhidul Mahbub, Ahamed Raihan, Rapsan Amin Anonto, and M. J. Alam

Abstract In this article, p-i-n solar have been simulated with SCAPS-1D solar photovoltaic device simulator to observe the effect of anti-reflecting coating over the p-i-n solar device. The device was constructed using AlGaAs. GaAs was used as P^+ , P, i, n^+, n-layer with changing doping in it. It is observed in the proposed design that there is a significant impact of this layer in p-i-n PV Device. In the p^+-p-i-n-n^+ device, 36.948% efficiency has been observed with anti-reflective coating. But without anti-reflective coating, it went down only at 24.3850%. Almost 12% variation with a layer. In the device of p^+-p-i-n, 33.5399% efficiency has been observed with anti-reflective coating. In the device of p-i-n, 33.4215% efficiency has been observed with anti-reflective coating. But without anti-reflective coating, it went down only at 22.0723%. Almost 11% variation observed with a layer. So, it can be said that anti-reflecting coating reduces the loss and increases the efficiency of a p-i-n PV Device.

S. S. Mahtab (✉)
Department of EEE, Feni University, Feni, Chittagong Division, Bangladesh
e-mail: mahtabshahzad@gmail.com

R. Akter
Department of EEE, Mymensingh Engineering College, Mymensingh, Bangladesh
e-mail: rupa.mec.eee.bd@gmail.com

T. Mahbub
Department of EEE, Independent University of Bangladesh, Dhaka, Bangladesh
e-mail: tony.mahbub@gmail.com

A. Raihan
Department of CSE, Jahangirnagar University, Savar, Dhaka, Bangladesh
e-mail: araihan13@gmail.com

R. A. Anonto
Department of EEE, American International University Bangladesh, Feni, Bangladesh
e-mail: rapsanaminanonto@gmail.com

M. J. Alam
Department of EEE, Feni University, Feni, Bangladesh
e-mail: alameee1993@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_20

Keywords P-i-N PV device · GaAs · 3–5 material · Photovoltaic · ARC · Anti-reflection coating

1 Introduction

The global average electricity consumption in 2006 was approximately 1610^{13} W [1]. The cumulative incident of solar radiation at ocean level is approximately 210^{17} W per day, which is more than 12,500 times the centre global energy utilization in 2006 [2]. It demonstrates the amazing potential solar power for providing the planet with electricity [3]. Photovoltaic devices are one way to harvest the sun's power [4]. The big field of delivering this enormous amount of energy is a fossil fuel with scarce resources [5]. A few of these capabilities are accessible only in a very few areas around the world; security also poses a challenge to the stability and sovereignty of national and regional entities, as well as global security [6, 7]. The protection issue and cost of disposing of radioactive material and hazardous waste renders the use of nuclear and chemical technology a dubious option. The fossil fuel is also expensive for transmission in remote zones. Accordingly, the interest in renewable energy has risen the last few decades along with the world's energy requirements, the alarms of global warming, and oil prices [8, 9]. A PV device is a solid-state semiconductor system that, when stimulated by photons, produces electricity from DC (direct current) [10].

As photons come into contact with the device's atomic structure, they dislodge electrons from the atoms. Which leaves a vacuum that makes other free electrons usable. If the device fabricates a PN junction, the dislodged photons flow towards the junction's P side. The effect of this electron transfer is a flow of electrical current which can be redirected through electrical contacts from the exterior of the membrane to generate energy [11, 12]. PV device's conversion η is measured as the ratio of input power(radiant energy) to power output (electric energy). In a p-i-n solar device, an intrinsic region is inserted between p- and n-type layer of a GaAs p-n solar device [13, 14]. In this case, J_{sc} is greater than the conventional p-n solar device. Sue to the additional carrier contribution of the intrinsic region. As a result, efficiency increases. In quantum well PV devices, variation among the appearance of photo energy levels to sustain photons from different energy sources is the incidence energy of the solar spectrum and single strip width. There is a fundamental similarity in this case, as seen is the main issue is the generation of multiple light-generated quasi-Fermi rates. However, the carriers' transport between the two approaches is considerably different. In quantum well solar device approaches, transportation is achieved by letting the carriers escape through light absorption at each localized energy level. The time of escape should be faster than the time of recombination to maintain a high collection η [15]. Viability of the escape mechanism is well demonstrated by infrared photodetectors with a large range of intra-sub-band detectors. Localized band solutions have another benefit in that successive localized energy rates may have different energies, thus allowing for a large number of successful band gaps and high efficiencies [9, 12, 16, 17].

The reflectivity, R(E), is significantly reduced by using an anti-reflective coating, and losses of reflection can be as small as possible [18]. In this article, the solar devices with anti-reflective coatings and without anti-reflective coatings are modeled on a reflectivity equal to zero for all wavelengths [10]. The impact of anti-reflective coatings has been checked in PIN PV Device [19].

2 Device Structure

Practical solar cells have a perplexing formation with different layers besides to the p-and n-layer got in the facile p-n photovoltaic solar cell [20]. The motive of these layers is to assuage back and front periphery recombination and periphery reflection. A plan of a p-i-n relevance cell is represented in Fig. 1 indicating all the layers (window, p^+, p, i, n, and n^+) comprised to gain a high efficient PV solar cell.

Aside from p- also n-surface, it is been found the easiest p-n PV device, practical PV devices have a complex design with many surfaces [21]. These layers aim to reduce the recombination of the front and back surface and the reflections on the

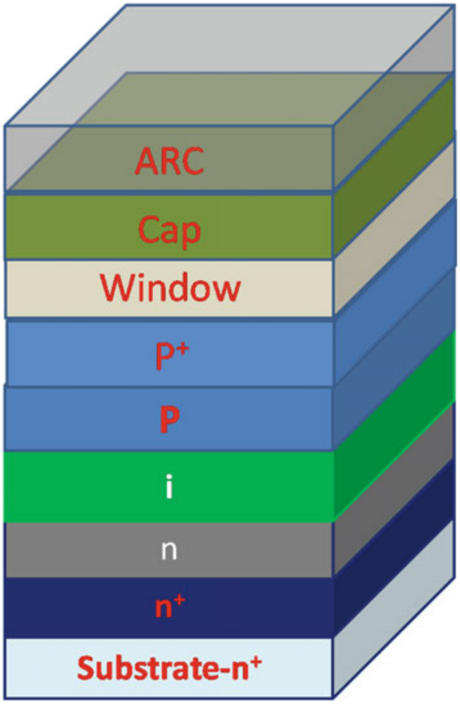


Fig. 1 Reference device p-i-n’s structure

Table 1 Reference device's materials

Layers	Materials
Window	AlGaAs
p^+	Gallium Arsenide
P	Gallium Arsenide
İ	Gallium Arsenide
N	Gallium Arsenide
n^+	Gallium Arsenide

surface. A reference device with p-i-n is given in Table 1, for all the levels (anti-reflective surface(ARC), window, p^+, p, i, n, and n^+) included obtaining a great η PV device. It includes of an anti-reflective surface, a cap-, window-, p^+, -, p-, i-, n- and n^+ surface mounted above the surface.

Various materials on the table have been used below in the reference device.

3 Simulation Results

In the simulation model, the below method is considered for calculating the system parameters.

Efficient band gap E_g. The made-in capacity determined by exercising the dominance carrier quantities on the interfaces *n* and *p*, the useful state solidities in the valence band N_{valen} and the conduction band N_{cond}. It is possible to write the dynamic equation for the valence band and the conduction band conducting unit geometry as follows:

$$E_{\text{cond}} = -\chi - q\varphi \tag{1}$$

$$E_{\text{valen}} = -\chi - E_g - q\varphi \tag{2}$$

The Concentration of minority carriers can also be estimated using

$$n_l = N_c \exp\left(\frac{F_n - E_{\text{cond}}}{kT}\right) \tag{3}$$

where, in *n*-side F_n is the Fermi equilibrium stage.

The charge carrier conductance was determined exercising the bi-polar formulas of drift-diffusion as given bellow: for electron,

$$J_n = q\mu_e n_f \frac{\partial E_{\text{cond}}}{\partial x} + qD_n \frac{\partial n_f}{\partial x} \tag{4}$$

And also for holes,

$$J_p = q\mu_h p_f \frac{\partial E_{\text{valen}}}{\partial x} + qD_p \frac{\partial p_f}{\partial x} \tag{5}$$

For the measurement of carrier trapping, the Shockley-Read-Hall (SRH) recombination principle was used [9]. The efficiency of the unit was theoretically determined by using the formula

$$\text{FF} = \frac{P\text{max}}{I_{\text{sc}} \times V_{\text{oc}}} \tag{6}$$

$$\eta = \frac{I_{\text{sc}} \times V_{\text{oc}} \times \text{FF}}{P_{\text{light}}} \tag{7}$$

where FF is considered as fill factor, P_{max} is known as maximum power, V_{oc} is as open-circuit voltage, Isc is known as short circuit current. The standards were accepted from the curve of I-V characteristics. The plight is the input power of 100 mw/cm^2 beneath 1 illumination with sunlight intensity at 1.5 AM.

The photocurrent produced at the short circuit by a solar device under illumination is dependent on the light from the incident. To relate the density of the photocurrent, J_{sc}, to the spectrum of incidents, the quantum efficiency of the device is needed (QE). $QE(E)$ is the probability of one electron being transmitted to the external circuit by an incident photon of power E. The photocurrent can then be written in $QE(E)$

$$J_{\text{light}} = q \int F(E) QE(E) dE \tag{8}$$

The dark current density J_{dark} (V) varies for an ideal diode as

$$J_{\text{dark}}(\text{V}) = J_0\left(e^{qV/k_B T} - 1\right) \tag{9}$$

The sign convention in photovoltaics for current and voltage is such that the photocurrent is positive. That is the reverse of the standard for electronic devices commonly used. The net current density in the device is with this sign convention

$$J(V) = J_{\text{SC}} - J_{\text{dark}}(V) \tag{10}$$

Total current can be express as

$$J = J_{\text{light}} - J_{\text{dark}} - \frac{V}{AR_{\text{sh}}} \tag{11}$$

where,

V is the terminal voltage and A is the area of the device and R_{sh} is the shunt resistance.

And voltage can be calculated as

$$\text{Voltage} = V - J\, R_s A \tag{12}$$

where R_s is the series resistance.

3.1 SCAPS-1D as the Simulation Tool

For this work, SCAPS-1D has been used as a simulation tool.A solar cell capacitance simulator (SCAPS) is a one-dimensional solar cell simulation programme developed at the Department of Electronics and Information Systems (ELIS) of the University of Gent, Belgium [22]. Several researchers have contributed to its development: Alex Niemegeers, Marc Burgelman, Koen Decock, Stefaan Degrave, Johan Verschraegen. The original programme is developed for cell structures of the CuInSe2 and the CdTe family [5]. Recent developments make the programme now also, applicable to crystalline solar cells (Si and GaAs family) and amorphous cells (a-Si and micromorphous Si) [9]

4 Results and Discussion

The maximum efficiency and current density of the short circuit were calculated for devices with different thickness of the window layer when other parameters, like the thickness of the other surface, were kept constant. To observe the effect of the anti-reflective coating simulation performed for with or without ARC. Both data are included in Table 2 for comparison between AlGaAs window layer system and GaP window layer system.

Table 2 Simulation results for anti-reflective coating, p^+ and n^+ layer

Anti-reflective coating (ARC)	Device layers	AlGaAs as the window layer		
		V_{oc} V	J_{sc} A/m^2	Max efficiency (%)
1	p^+-p-i-n-n^+	1.0054	445.1377	36.9481
0	p^+-p-i-n-n^+	0.9891	300.0228	24.3850
1	p^+-p-i-n	0.9880	407.1663	33.5399
1	p-i-n	0.9877	405.8243	33.4215
0	p-i-n	0.9734	273.5256	22.0723

ARC = 1 means present of anti-reflective coating, ARC = 0 means absent of anti-reflective coating

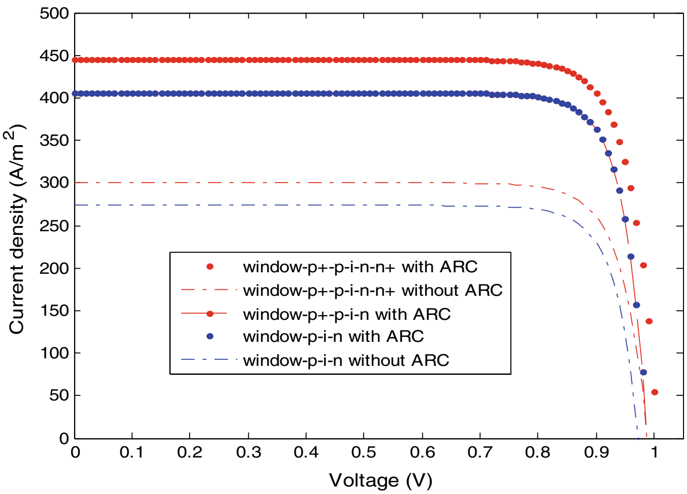


Fig. 2 V-I signifier of solar p-i-n devices with different extra layers

From this table, it is quite clear that there is a significant impact of this layer in p-i-n PV device. In the device of p^+-p-i-n-n^+, observed 36.948% efficiency with an anti-reflective coating. But without anti-reflective coating, it went down only to 24.3850% with 12% variation. In the device of p^+-p-i-n, 33.5399% efficiency observed with anti-reflective coating. The p-i-n device has 33.4215% efficiency with anti-reflective coating. But without anti-reflective coating, it shows 22.0723%, almost 11% varies with a layer.

Various curves are obtained from the simulation that is given below.

It can be observed from the curve using anti-reflective coating efficiency increase. This is due to minimize the reflection loss by ARC. It can also be observed from Figs. 2 and 3, that short circuit current and quantum efficiency is more for p^+-p-i-n-n^+ solar device than simple p-i-n PV device because both p^+ and n^+ layer reduce front and back surface recombination, respectively, so efficiency increase.

5 Conclusion

The p-i-n solar with SCAPS-1D solar photovoltaic device simulator is simulated here to observe the effect anti-reflecting coating over p-i-n solar device. The device was constructed using AlGaAs. GaAs was used as P^+, P, i, n^+, n-layer with dope changing. In p^+-p-i-n-n^+ an efficiency of 36.948% is observed in the anti-reflective coating and got 24.3850% without anti-reflective coating with 12% variation. In p^+-p-i-n observed 33.5399% efficiency with anti-reflective coating. In the p-i-n device, 33.4215% efficiency observed with an anti-reflective coating and 22.0723% without

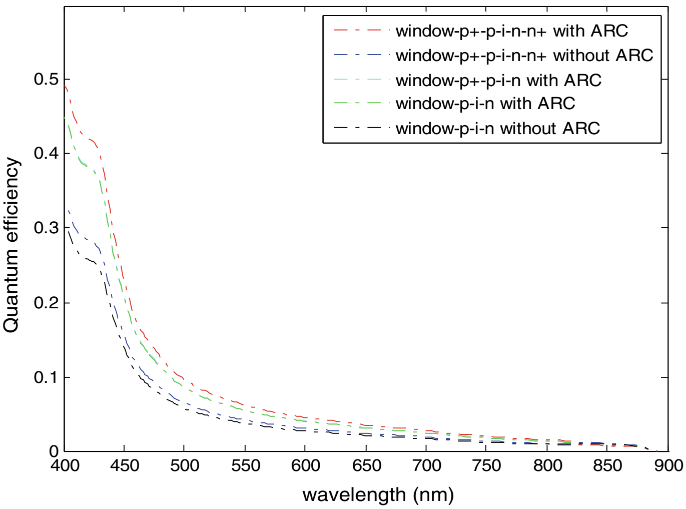


Fig. 3 Comparative analysis of p-i-n solar device's quantum performance with various additional layers

anti-reflective coating, almost 11% varies with a layer. So, it can be said that anti-reflecting coating reduces the loss and increases the η of a p-i-n PV device.

6 Limitation and Future Scope of This Study

Due to the lackings of fabrication instrument, this result can only be simulated. It is needed to fabricate in real life and check the validity if it matches with the simulation result or not. It is needed to check more materials for future increases in efficiency in term stability in solar energy. Lots more other 3–5 semiconductor materials are available like InSb and others which may give more efficiency with stability. Looking forward to checking them.

References

1. International Energy Outlook (2009) Energy Information Administration Office of Integrated Analysis and Forecasting U.S. Department of Energy. Retrieved online on 2009–06. http://www.jointsolarpanel.nl/fileadmin/jointsolarpanel/user/documents/seminar2004/stanleyzc04.pdf
2. Nelson J (2007) The physics of solar devices. Imperial College Press, London
3. Klausmeier-Brown ME, DeMoulin PD, Chuang HL, Lundstrom MS, Melloch MR, Tobin SP (1988) Influence of bandgap narrowing effects in p^+-GaAs on solar device performance. Confer Record Twentieth IEEE 1:503

4. Luque A, Hegedus S (2003) Handbook of photovoltaic science and engineering. Wiley, Hoboken
5. Kumakura K, Makimoto T, Kobayashi N (2000) Activation energy and electrical activity of Mg in Mg-doped $InxGa_{1-x}N$ ($x < 0.2$). Jpn J Appl Phys 39:L337
6. Emon IS, Ahmed SS, Milu SA, Mahtab SS (2019) Sentiment analysis of bengali online reviews written with english letter using machine learning approaches. In: Proceedings of the 6th international conference on networking, systems and security (NSysS ’19). Association for Computing Machinery, New York, NY, USA, pp 109–115. doi: https://doi.org/10.1145/3362966.3362977
7. Milu SA et al (2020) Sentiment analysis of bengali reviews for data and knowledge engineering: a Bengali language processing approach. In: Bindhu V, Chen J, Tavares J (eds) International conference on communication, computing and electronics systems. Lecture Notes in Electrical Engineering, vol 637. Springer, Singapore. https://doi.org/10.1007/978-981-15-2612-1_8
8. Twidell J, Weir T (2006) Renewable energy resources. Taylor and Francis, Milton Park
9. Chakma R, Mahtab SS (2019) Navigation and tracking of AGV in ware house via wireless sensor network. In: 2019 IEEE 3rd international electrical and energy conference (CIEEC), Beijing, China, 2019, pp 1686–1690. https://doi.org/10.1109/cieec47146.2019.cieec-2019589
10. Trondheim, Modeling of intermediate band solar Devices. Department of Physics, Faculty of Natural Sciences and Technology, Norwegian University of Science and Technology, Norway
11. Smith LL, Davis RF, Kim MJ, Carpenter RW, Huang Y (1997) Microstructure, electrical properties, and thermal stability of Au-based Ohmic contacts to p-GaN. J Mater Res 12:2249
12. Khan MFS, Mahtab SS (2019) PLC based energy-efficient home automation system with smart task scheduling. In: 2019 IEEE sustainable power and energy conference (iSPEC), Beijing, China, 2019, pp 35–38. https://doi.org/10.1109/ispec48194.2019.8975223
13. Greenpeace and European Photovoltaic Industry Association (2008) Solar Generation V-2008, Solar electricity for one billion people and two million jobs by 2020
14. Vassilevski KV, Rastegaeva MG, Babanin AI, Nikitina IP, Dmitriev VA (1997) Ti/Ni Ohmic contacts to n-type GaN. Mater Sci Eng, B B43:292
15. Wen TC, Lee WI, Sheu JK, Chi GC (2001) Characterization of p-type InxGa1-xN grown by metalorganic chemical vapor deposition. Solid State Elect 45:427
16. Mahtab SS, Alam MJ (2020) Numerical modeling and simulation of high-efficiency thin Cu(In, Ga)Se photovoltaic by WxAMPS. In: Bindhu V, Chen J, Tavares J (eds) In: International conference on communication, computing and electronics systems. Lecture Notes in Electrical Engineering, vol 637. Springer, Singapore. https://doi.org/10.1007/978-981-15-2612-1_40
17. Mahtab SS, Monsur A, Ahmed SS, Chakma R, Alam MJ (2018) Design and optimization of perovskite solar cell with thin ZnO insulator layer as electron transport. In: 2018 international conference on advancement in electrical and electronic engineering (ICAEEE), Gazipur, Bangladesh, pp 1–4. https://doi.org/10.1109/icaeee.2018.8643012
18. Ahmed SS, Milu SA, Emon IS, Mahtab SS (2020) Opinion mining of bengali review written with english character using machine learning approaches. In: Bindhu V, Chen J, Tavares J (eds) International conference on communication, computing and electronics systems. Lecture Notes in Electrical Engineering, vol 637. Springer, Singapore. https://doi.org/10.1007/978-981-15-2612-1_5
19. Strite S, Morkoç H (1992) GaN, AlN, and InN: a review. J Vac Sci Technol B 10:1237
20. Kumakura K, Makimoto T, Kobayashi N (2000) High hole concentrations in Mgdoped InGaN grown by MOVPE. J Cryst Growth 221:267
21. Mori T, Kozawa T, Ohwaki T, Taga Y, Nagai S, Yamasaki S, Asami S, Shibata N, Koike M (1996) Schottky barriers and contact resistances on p-type GaN. Appl Phys Lett 69:3537
22. Yamasaki S, Asami A, Shibata N, Koike M, Manabe K, Tanaka T, Amano H, Akasaki I (1995) P-type conduction in Mg-doped Ga0.91In0.09 N grown by metalorganic vapor phase epitaxy. Appl Phys Lett 66:1112

Microstrip Antenna with SVASTU Slot for UWB Applications

Shivani Chourasia, Sudhir Kumar Sharma, and Pankaj Goswami

Abstract The svastu slot stuffed with a rectangle microstrip antenna (MA) through a microstrip line and examines its realization. It is also called as a printed antenna. It comprises of the four-sided mark with svastu form which is assembled with one postern of the substrate ($30 \times 35 \times 1.6$ mm^3) and restricted base plane (16×8 mm^2) on the other postern. The FR-4 sealant substrate is used, and MA is delineated through the HFSS simulation technique which helps in providing support to give rise high frequency of 10.5 GHz which is working with the rate of occurrences of 4.1–14.6 GHz with VSWR < 2, and whereas the features procreate the delineate MA which is advisable for multiple UWB applications.

Keywords High frequency structure simulator (HFSS) · Ultra-wideband (USB) · Svastu slot

1 Introduction

MA has manifold benefits like lightweight, popularly priced, lean profile, informal to outline area whereas it can cover entirely larger technologies such as aircraft, satellite as well as wireless conversation [1, 2]. The major drawback of this MA is precarious transmission capacity. Whereas several efforts have taken to improve and raise the transmission capacity of the MA, although there have been some procedures to implement the transmission capacity with the help of parasitic patches one of the two same different layers. The similar layer in case parasitical patches are showing on some dissimilar layer then the solidity of the antenna raises or if the particular is nonce on a similar layer, then the lateral magnitude of the antenna raises. In harmony

---

S. Chourasia (✉) · S. K. Sharma
Department of Electronics and Communication Engineering, Jaipur National University, Jaipur 116067, India
e-mail: shivanicp01@gmail.com

P. Goswami
Department of Electronics and Communication Engineering, Teerthankar Mahaveer University, Moradabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_21

to overcome this obstacle to evolve excellent layer patch transmission capacity MA [3].

UWB, a wireless communication (WC) automation, which holds the attention of massive transmission capacity of a minimum 20% of center rate of occurrence government correspondence commission, designated a transmission capacity of 7.5 GHz, i.e., in distinction to 3.1–10.6 GHz during UWB tolls [4, 5].

UWB engender by short span throbbing in the magnitude of picoseconds so radical wide cord will give very elevated data quota up to multiple hundred Mbps and aphoristic period throbbing also eschew multipath evanescent. UWB articulation systems have newly acknowledged great thinking in the cellular universe [6]. This a broadly used high tech in scanning system and frontier sensing tools. This transmission capacity satiates the system needs for S-DMB, WLAN as well as CMMB. UWB antenna is having less return loss ($S11 < -10$ dB), high emission productiveness over UWB against 3.1–10.6 GHz [7].

One of the transcendent demanding roles is to expand the UWB antenna for adaptable tools as well as the technology which gives a very trifling portion for antenna settling. In a compact four-sided MA is delineating where U pattern slot has been taken to develop the impedance transmission capacity. Also, it is a requirement of an antenna that has several bands as well as it exhilarates exigency for UWB tools.

The svastu slot brimming four-sided MA is delineating prosperously and outcome after facsimile fulfill transmission capacity needs for UWB tools. In this depiction, for a rate of occurrences range of 4.1–14.6 GHz, using this range the $\mathrm{VSWR} < 2$, and $S11 < -10$ dB is realized. Its E- as well as H-sphere radiation molds are balanced over this rate of occurrences range [8, 9].

2 Microstrip Antenna

In the time of 1970s, different techniques have become excellent from over all the existence where the idea of MA is distinguished in the year 1953 and all the legitimate techniques like patents have been distributed in the time of 1955. It typically examines and comprises a lean strip that has been fixed on the head of the ground plane [10] which is depicted in Fig. 1.

The strip just as the ground sphere is separated by a dielectric membrane (intimate to as the substrate). There are different and various assorted substrates that can be taken into consideration for the outline of MA. There are distinct substrates that utilize the layout of MA, and their dialectic constraints are for the most part in the scope of $2.2 \leq \in r \leq 12$. The most beneficial is that these generally satisfy the receiving wire performance and are large substrates whose dielectric steady is in the optional finish of the range which is flexible in nature. Since they provide extremely bigger transmission limits and approximately headed fields for outflow into space, however, at the danger of bigger part [11, 12].

The transmission strip can be in any shape like might be square, four-sided, lean strip (dipole), round shape, egg molded, trilateral, or some other structure [13]. In

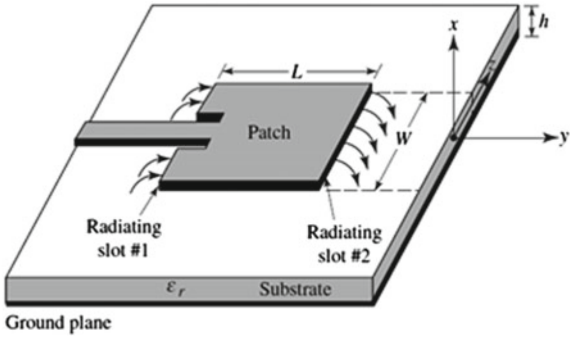


Fig. 1 Block diagram of microstrip antenna

the event that the more important demanding of MA, they are utilized in numerous advances where they are intensely required, for example, airplane, rocket, satellite, rocket instruments, where size, weight, cost, execution, simplicity of establishment, just as streamlined profile are limitation in the current situation, there are enormous quantities of government and business devices correspondingly portable radio just as remote innovation that have same necessity. To coordinate a similar degree of the requests, MA is comprehensively utilized and demanded [14–16].

3 Antenna Design

MA is delineated with the help of ANSOFT HFSS that is the highest rate of occurrence formative simulator. This is for forming MA with very limited and restricted components that have been used four-sided patched and the thickness of 11.964 mm as well as the length of 16 mm (Table 1).

The substrate is formed of FR-4 epoxy stuff having a diameter of 30 mm, length 35 mm, and compactness 1.6 mm. FR-4 epoxy is a dispensable nonconductor that is inexpensive as well as accessible. Which is having a stooge situation is 2.5 mm aside from situation which overlays an integrated ultra-wideband (Fig. 2).

Table 1 SVASTU dimension

S. No.	Parameter	Dimension	Material
1	Substrate	$W_{\text{sub}} = 30\,\text{mm}L_{\text{sub}} = 35\,\text{mm}H_{\text{sub}} = 1.6\,\text{mm}$	FR-4
2	Rectangular	$L_{\text{p}} = 16\,\text{mm}W_{\text{p}} = 11.964\,\text{mm}$	Copper
3	Ground plan	$W_{\text{g}} = 16\,\text{mm}L_{\text{g}} = 8\,\text{mm}$	Copper
4	Svastu slot	$6 * 1.5\,\text{mm}^2$ straight slot $L_{\text{S}} = 2.2\,\text{mm}$ $W_{\text{s}} = 1.5\,\text{mm}$	-
5	Feed line	$W_{\text{f}} = 3.01\,\text{mm}L_{\text{f}} = 8\,\text{mm}$	Copper

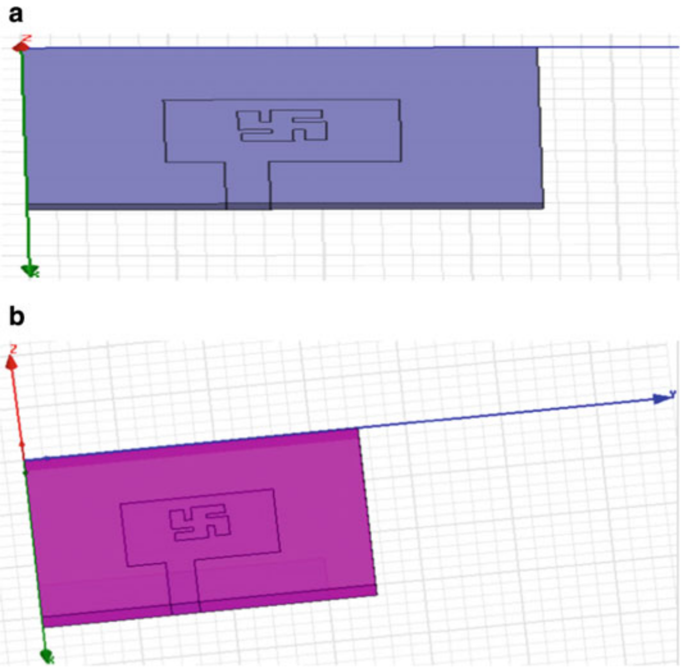


Fig. 2 **a** Svastu slot loaded MA. **b** Svastu slot loaded MA

4 Simulation Result

The outcome which is elucidating here is imitated on HFSS software. Circumstances that are used for overseeing substantial is inclination as skilled E furthermore for transmission box usual air reputation are fit which is elucidated in Fig. 3 which demonstrate reoccur loss vs. rate of occurrences plot.

Reoccur loss provides a quantity of power that has been reproduced separately input port. Whereas the UWB antenna is reoccurred detriment below −10 dB is treated to be precisely economical. Further, antenna layout reoccur detriment is lesser as compare −10 dB in the rate of occurrence limits 4.1–14.6 GHz.

The E-sphere is illustrated as the sphere which accommodates the electric field trajectory and pioneers the uttermost transmission meantime; the H-sphere is the sphere accommodating the magnetic field trajectory as well as the administration of altitude emission.

The x–z sphere upgrading sphere which is having some certain smooth angle φ is the principle E-sphere. Meantime for the x–y sphere, azimuth sphere, which is having a few specific altitude angle θ, is principle H-sphere and in Fig. 4 has elucidated 2-D *E*-sphere Radiation pattern (RP) at a unique managing rate of occurrence.

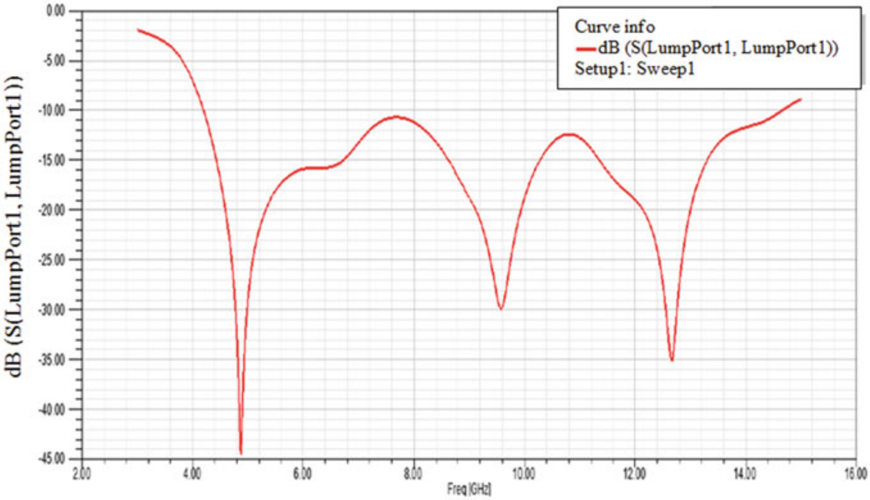


Fig. 3 Variation of Frequency vs Return loss (S11)

For the *x*–*y* sphere is a smooth plane with some specific levitation angle θ is criterion H-sphere which has been elucidated in Fig. 5, 2-D H-plane emission design at a dissimilar operating rate of occurrence (Fig. 6).

Figure 7 displays the voltage standing wave ratio (VSWR) plot as compared to the rate of occurrence in GHz. VSWR is directly the ratio of crest amplitude of eminence twirl to the littlest amplitude of eminence twirl. VSWR below 2 is thought hardy for an antenna. For this delineation, VSWR is lesser than 2 take-ups away 4.1–14.6 GHz.

5 Conclusion

A rectangular microstrip antenna is advisable for UWB tools that is delineation. It shows a transmission capacity of 112.31% (4.1–14.6 GHz, centralize at 9.35 GHz). The suggested outlet of this antenna could be used for a multiple of UWB tools in addition to greater speed data transfers, wireless connectivity in the middle of UWB-enabled equipment, and diversification of medical tools.

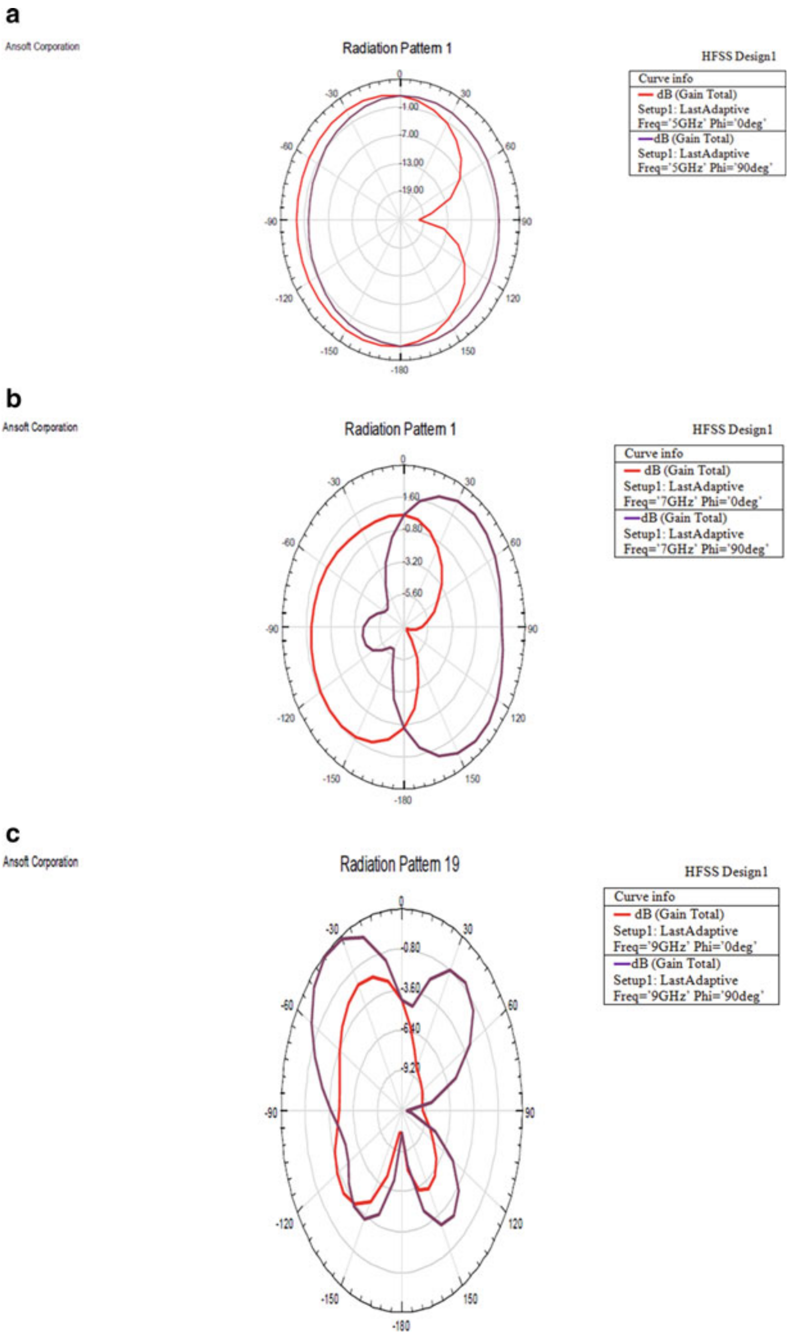


Fig. 4 **a** 2-D E-plane RP at 5 GHz. **b** 2-D E-plane RD at 7 GHz. **c** 2-D E-plane RD at 9 GHz. **d** 2-D E-plane RD at 12 GHz. **e** 2-D E-plane RD at 12 GHz

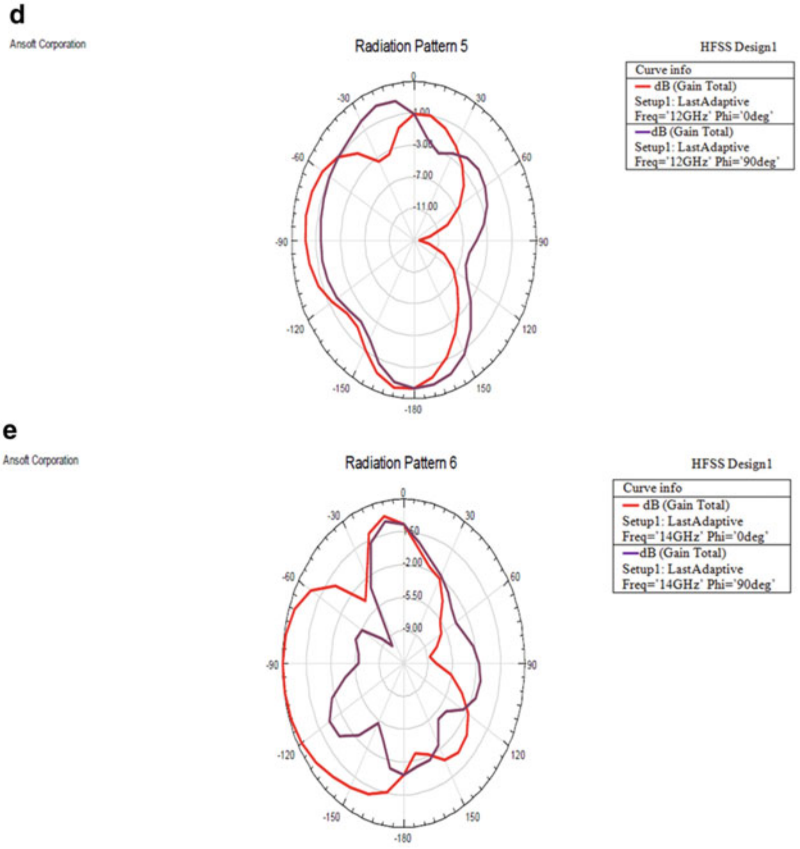


Fig. 4 (continued)

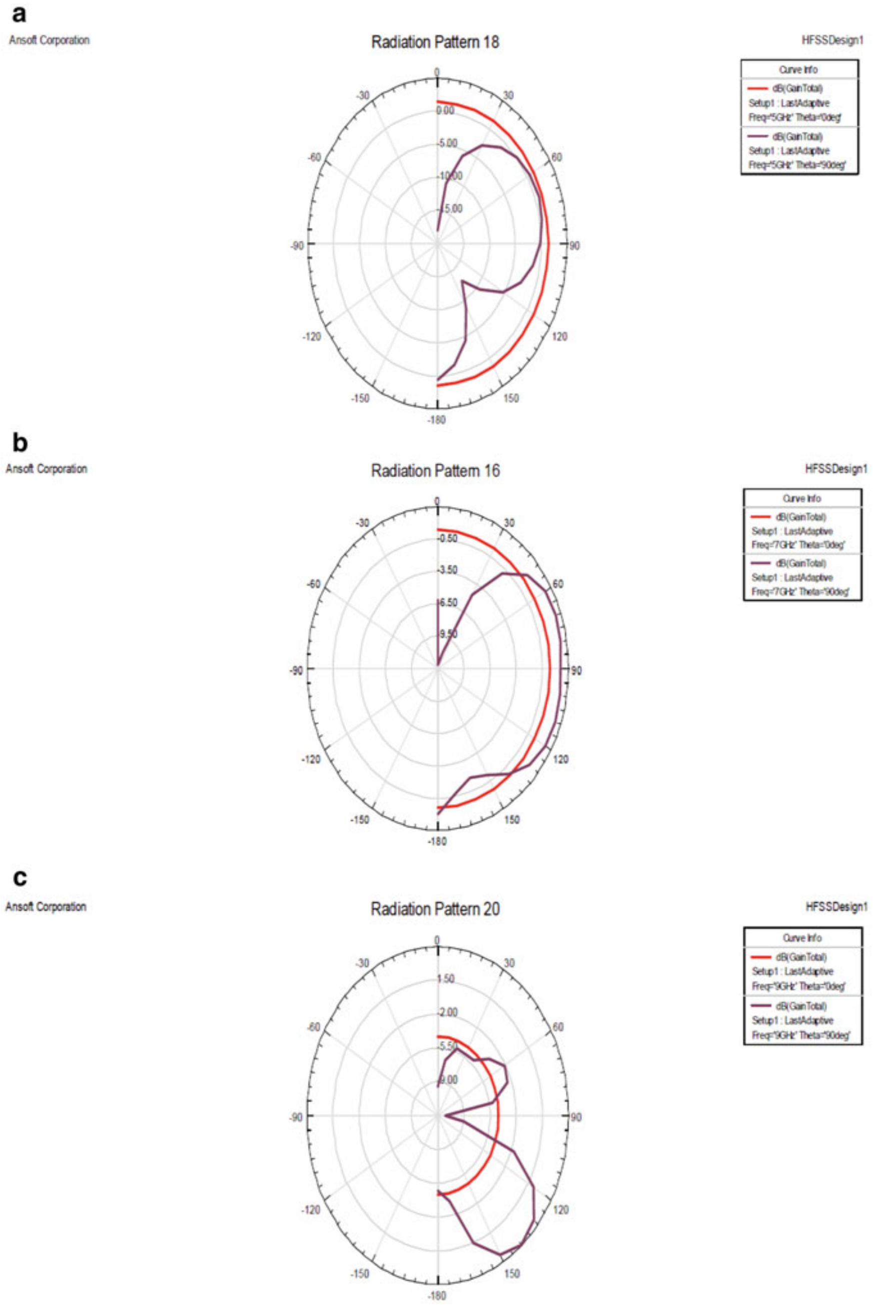


Fig. 5 **a** 2-D H-plane RD at 5 GHz. **b** 2-D H-plane RD at 7 GHz. **c** 2-D H-plane RD at 9 GHz. **d** 2-D H-plane RD at 12 GHz. **e** 2-D H-plane RD at 14 GHz

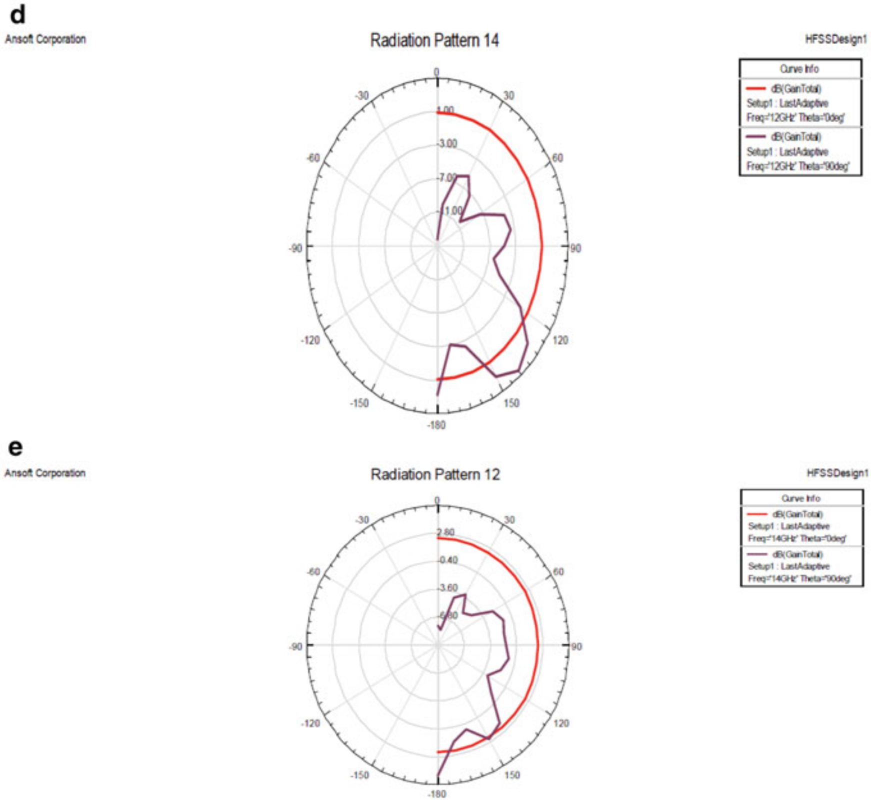


Fig. 5 (continued)

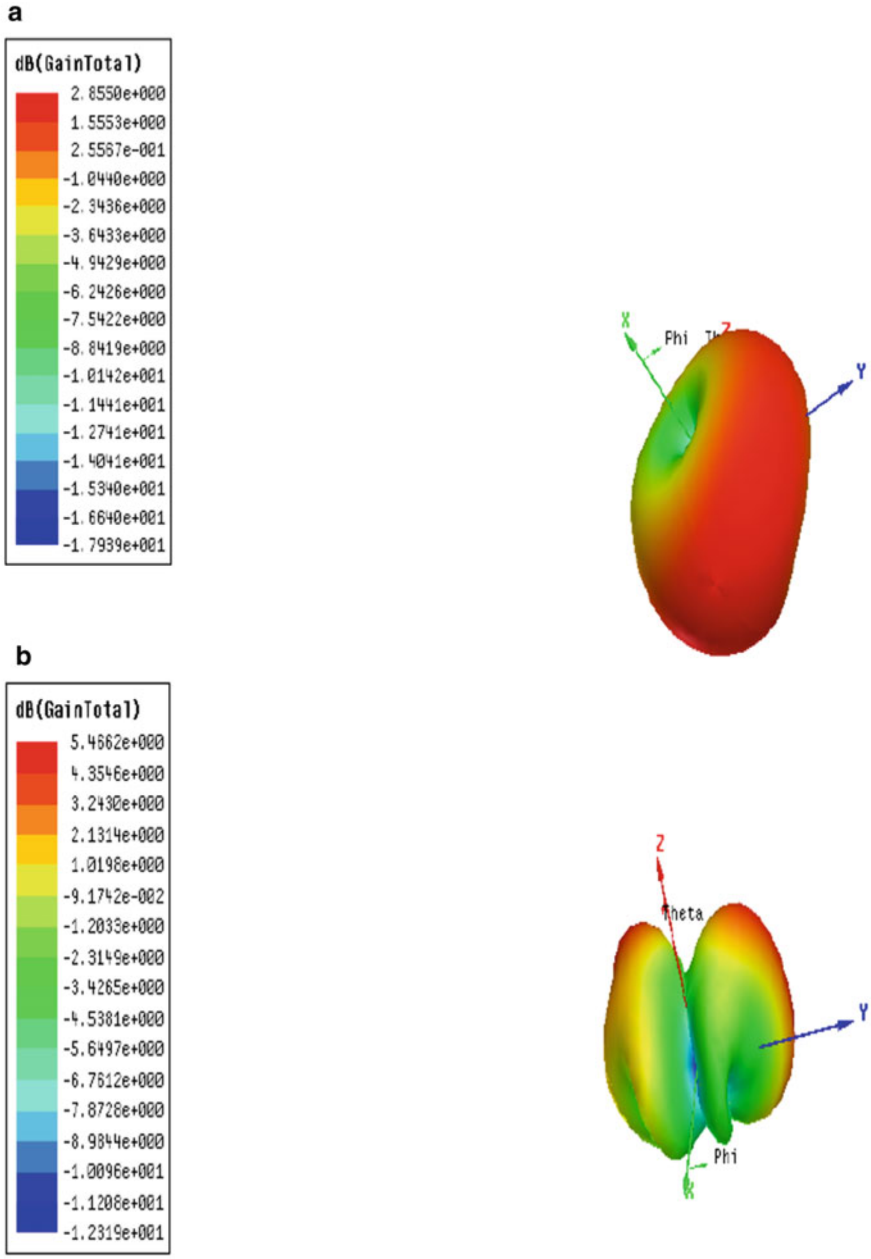


Fig. 6 **a**3-D RD at 5 GHz. **b** 3-D RD at 10 GHz

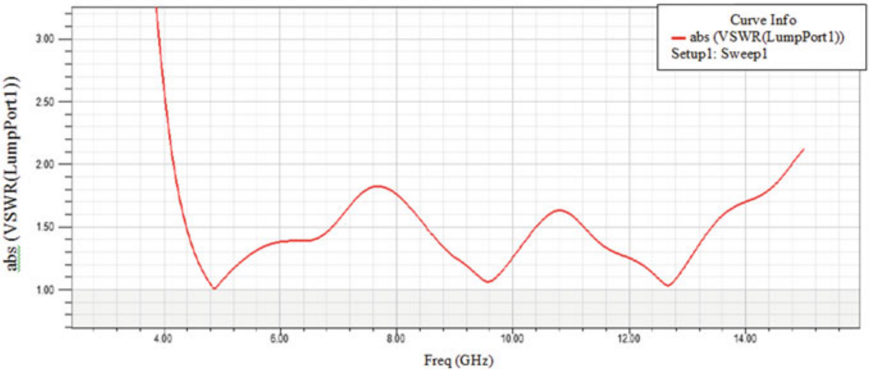


Fig. 7 Variation of Frequency vs VSWR

References

1. Sharma A, VishwakarmaRK (2014) Microstrip antenna with Swastik slot for UWB applications.In: 2014 IEEE students' conference on electrical, electronics and computer science, pp 1–5. IEEE
2. Prakasam V, Sandeep P (2017) Mode patterns in rectangular waveguide. Int J Trend Res Dev (IJTRD)
3. Luo J, Chi L, Li C, Sun B (2018) Side edge frame printed eight port dual band antenna array for 5G smart phone applications. IEEE Trans Antennas Propag 66(12):7412–7417
4. Sun L, Fang H, Li Y, Zhang Z (2018) Compact 5G mobile phone antennas with tightly arranged orthogonal mode pairs. IEEE Trans Antennas Propag 66(11):6364–6369
5. Porchan NO, Al- YIA, Ali AH (2019) Eight element dual polarized MIMO slot antenna system for 5G smart phone applications. IEEE Access 7:15612–15622
6. Ayyappan M, Chadran J (2016) Design and analysis of circular micro-strip antenna at 2.4 GHz with Fr-4 substrate. Int J Adv Res Electr Electron InstrumEng 5(4)
7. Prakasam V, Sandeep P (2018) Design and analysis of 2×2 circular micro-strip patch antenna array for 2.4 GHZ wireless communication applications. Int J InnovEng Manage Res 7(12)
8. Naik KK (2018) Asymmetric CPW-fed SRR patch antenna for WLAN/WiMAX applications. Int J Electron Commun (AEÜ) 93:103–108
9. Checkatla AR, Ashtankar S (2019) Compact microstrip antenna for 5G mobile phone application. Int J Appl Eng Res 14(2):108–111
10. Fallahpour M, Zoughi R (2018) Antenna miniaturization techniques: a review of topology- and material-based methods. Antennas Propag Magaz IEEE 60(1):38–50
11. Mohammad HA, Mohammad U, Chughtai T, Nasir J (2018) Cross polarized 2×2 UWB-MIMO antenna system for 5G wireless applications. PIER-M 76(5):157–166
12. Wang Y, Zhu L, Wang H, Yang G (2018) Design of compact wideband meandering loop antenna with a monopole feed for wireless applications. PIER Lett 73:1–8
13. Chen T, Li ZN (2019) Shared surface dual band antenna for 5G application. IEEE Trans Antennas Propag 1–1
14. Khan R, Abdullah A, Soh PJ (2018) User influence on mobile terminal antennas: a review of challenges and potential solutions for 5G antenna. IEEE Access 6:77695–77715
15. Haun C, Kuster Li N, Ofli E (2007) The effect of hand phantom on mobile phone antenna OTA performance. EuCAP Edinburgh
16. Jain S, Tomar PS, Tomar GS (2012) Design & analysis of proximity fed circular disk patch antenna. Int J Emerg Technol Adv Eng 2(10):126–131

Design and Implementation of a Smart Helmet System for Underground Miner's Safety

S. M. Minhajul Alam, Arnob Barua, Ahamed Raihan, M. J. Alam, Rocky Chakma, S. S. Mahtab, and Chitra Biswas

Abstract In industrial applications, safety is determined as a crucial theme. With time, many solutions are used to save labors from traumatic events or constructional failures. A wearable helmet is exhibited in this study, warns leaking toxic gas in a large amount. This study is also aid for petro-chemical to save the lives as well as the safety of workers. This prototype provides real-time monitoring of harmful gases, temperature, humidity, and worker's heart-rate. To overcome the hazardous situations, this system provides an emergency alarm for the monitoring station. With the aid of unified sensors, the helmet is adroit to execute nearby observing actions of the labor and rapidly forwards alarm regarding the un-avoidable collision. This system is very user-friendly and cost-effective.

Keywords Internet of things (IoT) · Radio frequency (RF) · GPS/GPRS

S. M. Minhajul Alam · A. Barua · R. Chakma · C. Biswas
Department of EEE, USTC, Chittagong, Bangladesh
e-mail: alameee1993@gmail.com

A. Barua
e-mail: arnob303@gmail.com

R. Chakma
e-mail: rocky.cht@gmail.com

C. Biswas
e-mail: chitra.biswas86@gmail.com

A. Raihan
Department of CSE, Jahangirnagar University, Savar, Dhaka, Bangladesh
e-mail: araihan13@gmail.com

M. J. Alam (✉) · S. S. Mahtab
Department of EEE, Feni University, Trunk Road, Feni 3900, Bangladesh
e-mail: alameee1993@gmail.com

S. S. Mahtab
e-mail: mahtabshahzad@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_22

1 Introduction

In this system, a safety helmet is made for the worker, rescuer, and miner. The natural condition of the current situation like humidity, temperature, heart-rate, and harmful gas detection can be detected by radio frequency (RF) communication. The position of a worker can be found using GPS/GPRS. The helmet's sensor senses the data from weather and the human body [6]. Then, this data is sent to the monitoring module by RF transmitter. After that, the monitoring section can receive all the data from the RF receiver and display.

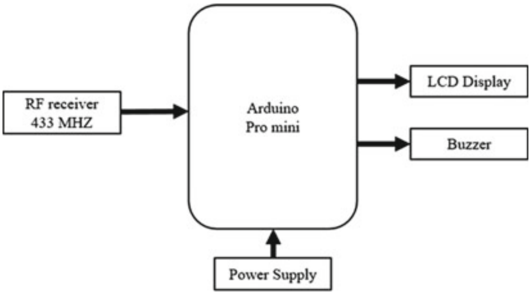
The demand for coal as energy resources is always important and significant. Nevertheless, thousands of people have forfeited their breath in mining accidents, all over the world. In their article, Jing change, Qinggui Cao and Yonjige Yang listed 100 of the major mining accident which had taken place from 2001 to 2010 [1]. As most of the coal mines in the North East region of the country are still in a primitive state, the mining accidents here are also very frequent. The main reason is these accidents occur due to the presence of methane and carbon monoxide gas in theses mines. These gasses are colorless, odorless and are undetectable by human sensors [2]. Such kind of accident can be controlled in the prophecy of the explosion process by executing microcontrollers and sensors and to develop a siren scheme before demanding climatic level. A steady supervising is important which further recommends some efficient and solid sensing method. To sense the existence poisonous gas, several techniques can be accommodated, using semiconductor type gas is much more effective among the techniques. The advantage of these sensors is, can be mounted in the coal mine location [3]. Apart from this, some disadvantages also can be calculated at the time of mining. The damage of the sensor apparatus from an accident often took place. Using robot can be another fruitful technique [4]. In no doubt, these kind of robots are good to use but for a country like India where industrialists are not much concerned about the safety of the workers, a robot cannot be imagined. However, considering cost-effectiveness and some other factors, safety helmet technology can be the further solution for the coal mine workers, whereas a smart safety helmet with updated sensor array will be presented to sense data and also a wireless modem will be available to transmit the data [5].

2 Literature Review

A helmet is an embodiment of defensive equipment on the head to shield it from wounds. In 900 BC, Assyrian soldiers used a helmet which is considered the oldest one, he wore bronze or thick leather helmets to shield the head from the unsharpened body and arrow strikes and sword blows in war. In civilian life, helmets are also used for enjoyment activities, transportation, and ports. On the other hand, different types of the innovative helmet can be seen in the modern world [7]. Among these, Baseball batting helmets is one which has a bolster safety over the ear and it saves the mouth and jaw from the wound. For safety from wind and rain, motorcycle helmets generally have flip-down face screens, and also a projecting visor is randomly used to shield the eyes from glare [8]. Advance technology can be seen in modern firefighter's helmets where include communication systems, masks, and other accessories which save the face and back of the head from accident, electricity, and fires. There is another helmet named mixed martial arts helmets where a pair of ear pads are included to restrict serious injuries to the players, those who do not commonly consume such force to the ears [9].

To upgrade safety, comfort, and current innovations for the target market, Brembo helmets are devised, these helmets are in the new varieties of automatic fit belt fastening of the helmet and the configuration of the visors. Due to hot weather and humid condition, forest workers will remove the head shields. But, it is essential for the workers to wear helmet during work which is claimed by occupational safety and health administration (OSHA) in consequence to face a compliance difficulty [10]. To identify which features, provide to forest labors' thermal embarrassment, this research assessed subjects' psychophysical and physiological reactions at the time of tasks estimating the assignment of forest labor in a large amount of temperature condition are observed throughout the summer in southeastern united states. In this study, three more helmets assessments were practiced: (a) a standard helmet, (b) an actively ventilated helmet, and (c) a passively ventilated helmet. It was observed that none of the physiological variables are examined for the body loaded with helmets [11]. Also, it is observed that dry bulb temperature and wet bulb temperature are differed in the tested helmets. psychophysical outcomes appears as air circulation will commits as a prominent suitable helmet and its fit, and weight are significant features in helmet pattern which is much important to factories and industries. Protective helmets for utilizing in humid and hot atmospheres must be altered to be convenient and motivating forest labors to wear and obey with OSHA regulations. 2001 Elsevier Science B.V. All rights set aside. In consequence, the head substantially regulates heat transfer and performs a crucial role in governing comprehensive individual comfort. Thermal comfort has suited progressively vital in designing different kinds of helmets (safety, firefighter, bicycle, etc.) as it can importantly ameliorate users' safety and health. Attaining thermal comfort implies considering multiplex airflow, moisture movement, and heat transfer [12]. Beginning evaluation of thermal properties and comfort of helmets disseminate the subsequent similarities between industrial safety helmets and cricket: the purposed of the these helmets is to save the user head [1] in case of shortest impact with a thing; both cricket players and industrial labors [2] are displayed to an utmost environmental situation for a prolonged spell of time; most of the users [3] noticed that most of the convention helmets have more weight, and it produces uncomfort to wear due to poor ventilation. International cricket environments have moderate airflow and to analyze the thermal distributions for helmets is accepted as <0.8 ms À1 which is tested under laboratory. Considerably, the helmet design used for cricket and motorcycle are much similar in construction and the appearance will control the air flow in the helmets [13]. Existing research models confines that position of vents increases the airflow and reduces the heat. Usage of dissimilar polymeric materials such as acrylonitrile butadiene styrene, Polycarbonates and polyethylene will be a suitable choice for industry helmets.

3 Hardware Block Diagram of Helmet



Two basic elements of system block diagrams are considered for this project. In addition, the receiver and transmitter block diagram are suggested for this approach.

3.1 Block Diagram of Transmitter

An electronic circuit with a switched-mode power supply (SMPS) transmuted power operating switching appliance that is turned on and off at high frequencies, and elements stored inside such as capacitors or inductors to transmit power during the switching apparatus is in its non-conduction condition. Switching power supplies have excessive efficiency and are extensively utilized in a difference of electronic appliance, incorporating computers, and other sensitive appliance demanding steady and effective power supply. A switched-mode power supply is also familiar as a switching-mode power supply (Figs. 1 and 2).

MQ2 gas sensor performs on 5 V DC and takes out approximately 800mW. It also can recognize LPG, alcohol, propane smoke, hydrogen, methane, and carbon monoxide concentrations everywhere from 200 to 10,000 ppm. When semiconductor particles such as tin dioxide are warmed in the air at overpriced temperature, oxygen is sucked up on the exterior. In scrubbed air, donor electrons in tin dioxide are engaged about oxygen which is sucked up on the exterior of the sensing substances. This obstructs electric current pass. In the appearance of condensing gases, the exterior density of adsorbed oxygen attenuates as it behaves with the condensing gases. Electrons are subsequently

delivered into the tin dioxide, authorizing current to pass openly direct to the sensor. In this project, a prominent network technology (GSM) is applied to utilize for

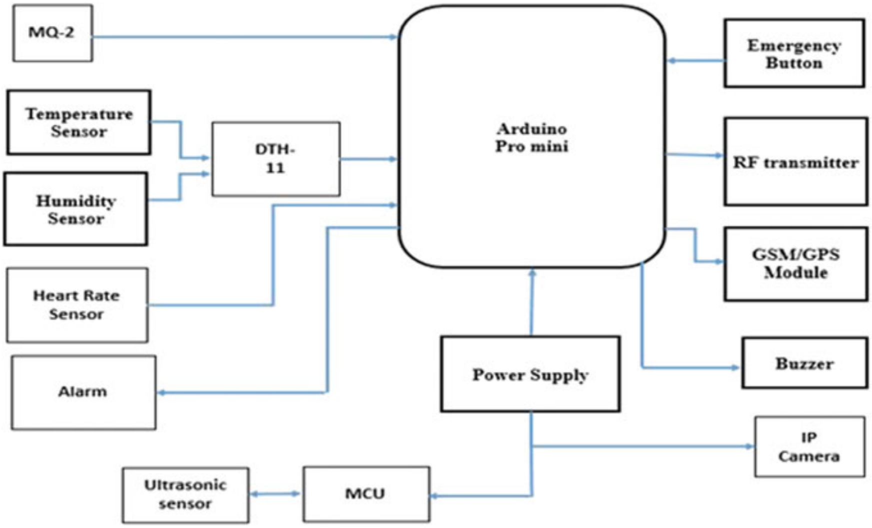


Fig. 1 Transmitter

Fig. 2 Receiver

transmission of SMS from sender to receiver. SMS sending and receiving is generally manipulated for omnipresent access of device and permitting breach control at here and there. The security alert network accords self-executing security observation. The system is efficient adequate to command user via SMS from a particular cell number to wide awake the circumstances of the industry machine as per to the user's demands and requirements. The second feature is that of security alert which is attained in a method that on the observation of intrusion, the system permits self-executing creation of SMS, therefore, warning the user against security risk [2].

4 Hardware Design

Here, a gas sensor, humidity, and temperature sensor, Arduino, and a buzzer have been used for emergency purpose. Our system includes an emergency sensing switch. The gas sensor is operated to recognize the weather bad conditions, the achievement is fed to the Arduino. The gas sensor and switch both are placed accurately in the helmet. Based on the gas sensors, humidity, and temperature sensor concentration gas sensors, and the humidity and temperature sensor provides an analog resistive result. Arduino unit, which controls all the functions of other blocks in this system. Indeed, Arduino accepts or reads data comes from the sensors and governs all the tasks of the entire system by utilizing these data. Arduino secures data from these sensors and it represents a digital data wireless tie in with to the output receiver edge (Figs. 3, 4 and 5).

4.1 System Result

When the worker presses the emergency button, the system will send the message to the monitoring section. A monitor will have a message and the worker's location (Latitude and Longitude) value, Google map includes this message. Then, the rescue team will find out and help the worker.

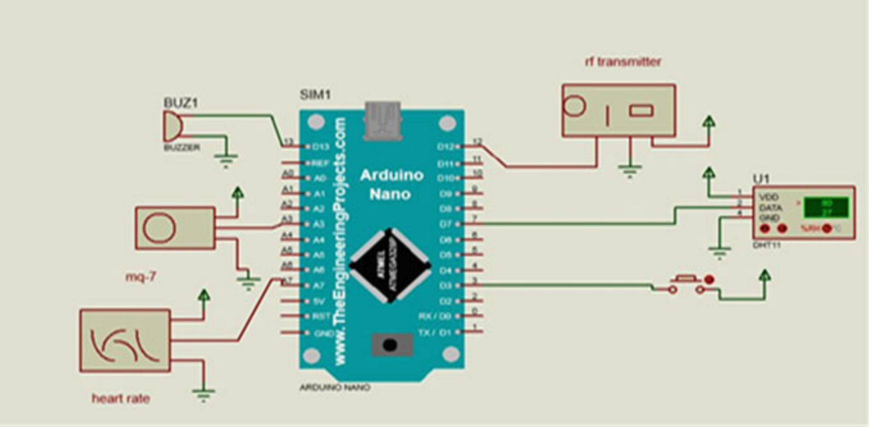


Fig. 3 Circuit diagram for sending

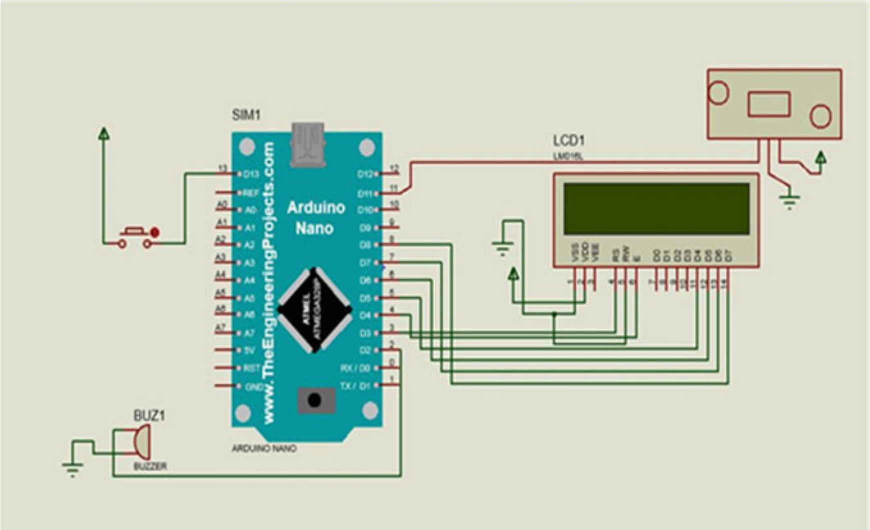


Fig. 4 Circuit diagram for receiving

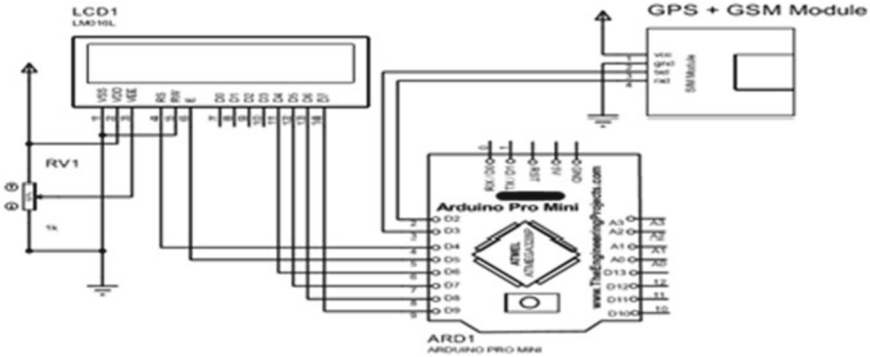
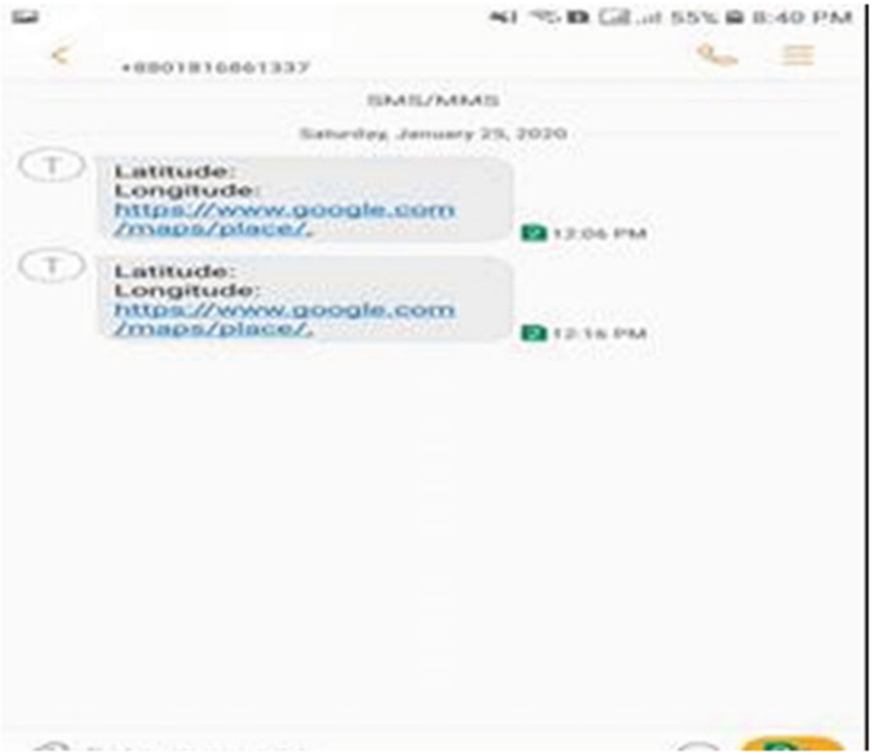


Fig. 5 Arduino accepts or reads data from Sensor

5 Result Analysis

This project gives an overview of the project introduction, overview, features, aims and objectives, scope, problem definition, background, and operation environment of the system. Here, the design (Fig. 6) and the circuit (Fig. 8) used to operate the system is shown below (Figs. 7 and 9).



+8801816661337

SMS/MMS

Saturday, January 25, 2020

Latitude:
Longitude:
https://www.google.com/maps/place/,

12:06 PM

Latitude:
Longitude:
https://www.google.com/maps/place/,

12:16 PM

Fig. 6 System send the message

6 Conclusion and Future Scope

A smart helmet band have been sufficiently devised in this project utilizing both smart technology GPS and GSM. The project led mandatory of put on a helmet to commence the flaming of a motor vehicle and interval of driving. If any unanticipated transformation in acceleration takes place afterwards, an accelerometer will supervise the transformation and a direct message with the position of the driver will be posted to the predetermined number utilizing GSM constituent. However, GPS/GSM constituent can't be able to address labor's particular position longitude and latitude value. This is a condition where a few solutions are discovered to the issue of expanded death percentage. In the future, the proposed system can connect the sensor's values to the online server with the internet of things technology.

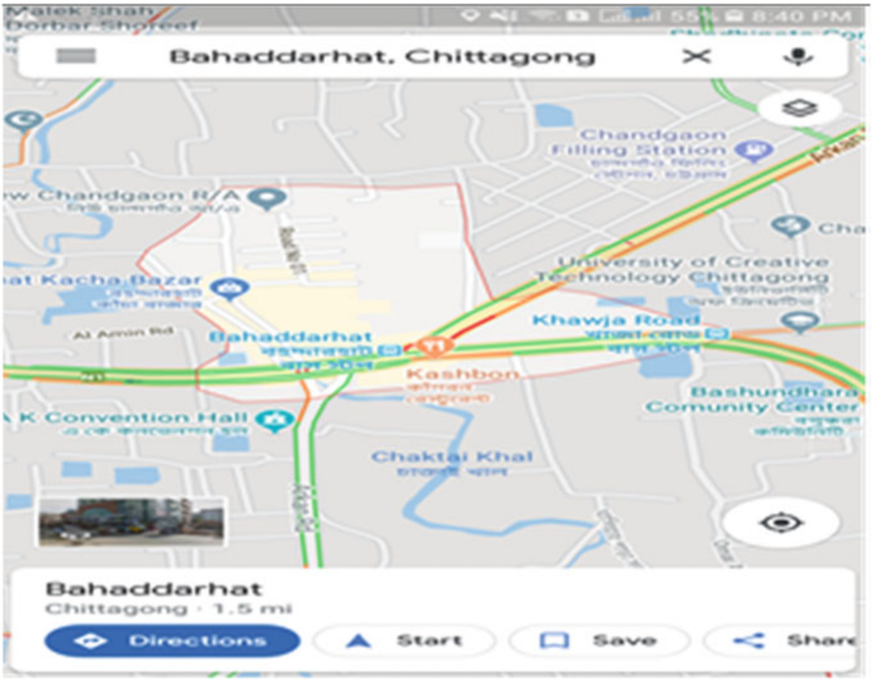
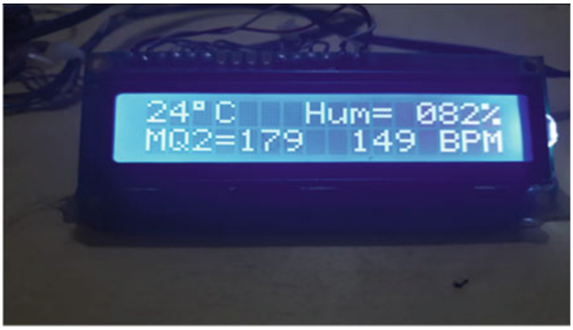


Fig. 7 Google map

Fig. 8 Sensor value



```
24°C    Hum= 082%
MQ2=179   149 BPM
```

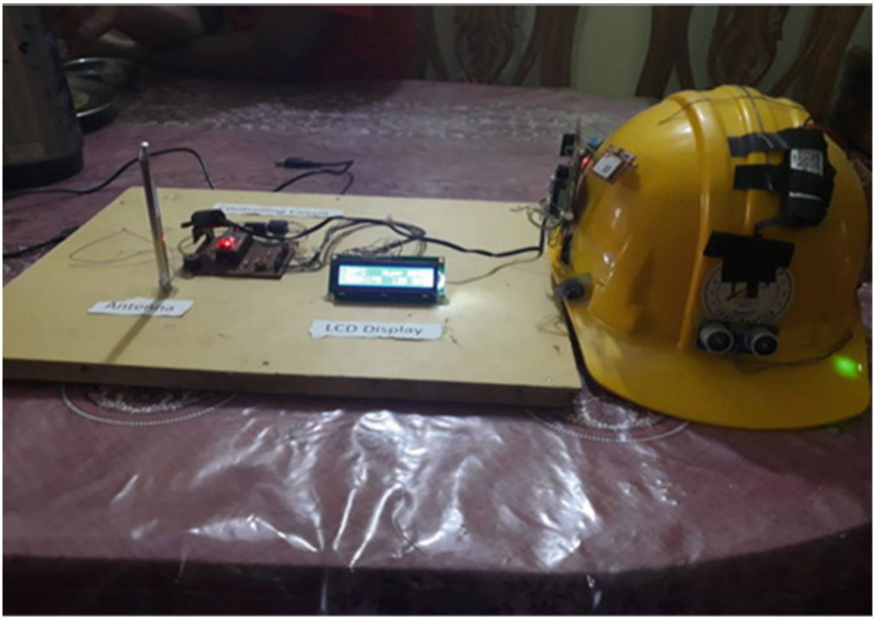


Fig. 9 Physical overview of the project

References

1. Khan MFS et al (2019) PLC based energy-efficient home automation system with smart task scheduling. In: IEEE sustainable power and energy conference (iSPEC), Beijing, China, pp 35–38
2. Chakma R et al (2019) Navigation and tracking of AGV in ware house via Wireless Sensor Network. 2019 IEEE 3rd international electrical and energy conference (CIEEC), Beijing, China, pp 1686–1690
3. Mahtab SS, Alam MJ, Khan AM, Uddm Z, Mamun AA, Uddin MM (2017)Optimization of InSb QWFET layer structure for high-speed and low power nano electronics applications. In: 2017 4th international conference on advances in electrical engineering (ICAEE), Dhaka, pp 707–711. IEEE. doi: https://doi.org/10.1109/ICAEE.2017.8255447
4. Mahtab SS, Monsur A, Ahmed SS, Chakma R, Alam MJ (2018) Design and optimization of perovskite solar cell with thin ZnO insulator layer as electron transport. In: 2018 international conference on advancement in electrical and electronic engineering (ICAEEE), Gazipur, Bangladesh, pp 1–4. IEEE. doi: https://doi.org/10.1109/ICAEEE.2018.8643012
5. Mahtab SS, Hossain MA (2019) Efficient and stable perovskite solar cell with TiO_2 thin insulator layer as electron transport. In: 2019 international conference on robotics, electrical and signal processing techniques (ICREST), Dhaka, Bangladesh, 2019, pp 54–58. IEEE. doi: https://doi.org/10.1109/ICREST.2019.8644093
6. Chakma R, Mahtab SS (2019) Navigation and tracking of AGV in ware house via wireless sensor network. In: 2019 IEEE 3rd international electrical and energy conference (CIEEC), Beijing, China, pp 1686–1690. IEEE. doi: https://doi.org/10.1109/CIEEC47146.2019.CIEEC-2019589
7. Khan MFS, Mahtab SS (2019) PLC based energy-efficient home automation system with smart task scheduling. In: 2019 IEEE sustainable power and energy conference (iSPEC), Beijing,

China, pp 35–38. IEEE. doi: https://doi.org/10.1109/iSPEC48194.2019.8975223

8. Emon IS, Ahmed SS, Milu SA, Mahtab SS (2019) Sentiment analysis of Bengali online reviews written with english letter using machine learning approaches. In: Proceedings of the 6th international conference on networking, systems and security (NSysS '19). Association for Computing Machinery, New York, NY, USA, pp 109–115. doi: https://doi.org/https://doi.org/10.1145/3362966.3362977
9. Milu SA et al (2020) Sentiment analysis of Bengali reviews for data and knowledge engineering: a Bengali language processing approach. In: Bindhu V, Chen J, Tavares J (eds) International conference on communication, computing and electronics systems. Lecture Notes in Electrical Engineering, vol 637. Springer, Singapore. https://doi.org/https://doi.org/10.1007/978-981-15-2612-1_8
10. Ahmed SS, Milu SA, Emon IS, Mahtab SS (2020) Opinion mining of Bengali review written with English character using machine learning approaches. In: Bindhu V, Chen J, Tavares J (eds) International conference on communication, computing and electronics systems. Lecture Notes in Electrical Engineering, vol 637. Springer, Singapore. https://doi.org/https://doi.org/10.1007/978-981-15-2612-1_5
11. Buurat J, Mahtab SS, Milu SA, Emon IS (2020) An automated Bengali text summarization technique using Lexicon based approach. In: International conference on innovations in computer science and engineering (ICICSE-2020). Springer "Lecture Notes in Networks and Systems"
12. Buurat J, Mahtab SS, Milu SA, Emon IS (2020) A pronoun replacement based special tagging system for Bengali language processing (BLP) In: International conference on innovations in computer science and engineering (ICICSE-2020). Springer "Lecture Notes in Networks and Systems"
13. Hoque F, Chakma R, Mahtab SS, Akter R, Ahmed SS. Design and developing real time interactive IIUC bus tracking system. J Innov Comput Sci Eng (JICSE) with ISSN 2278-0947 in vol 9(2), vol 10(1)

A Brief Review on Instance Selection Based on Condensed Nearest Neighbors for Data Classification Tasks

Yasmany Fernández-Fernández, Diego H. Peluffo-Ordóñez,
Ana C. Umaquinga-Criollo, Leandro L. Lorente-Leyva,
and Elia N. Cabrera-Alvarez

Abstract The condensed nearest neighbor (CNN) classifier is one of the techniques used and known to perform recognition tasks. It has also proven to be one of the most interesting algorithms in the field of data mining despite its simplicity. However, CNN suffers from several drawbacks, such as high storage requirements and low noise tolerance. One of the characteristics of CNN is that it focuses on the selection of prototypes, which consists of reducing the set of training data. One of the goals of CNN seeks to achieve the reduction of information in such a way that the reduced information can represent large amounts of data to exercise decision-making on them. This paper mentions some of the most recent contributions to CNN-based unsupervised algorithms in a review that builds on the mathematical principles of condensed methods.

Keywords Prototypes · Nearest neighbor algorithms · Classification

---

Y. Fernández-Fernández (✉)
Universidad Politécnica Estatal del Carchi, Tulcán, Ecuador
e-mail: yasmany.fernandez@upec.edu.ec

Y. Fernández-Fernández · D. H. Peluffo-Ordóñez · A. C. Umaquinga-Criollo ·
L. L. Lorente-Leyva
SDAS Research Group, Ibarra, Ecuador
e-mail: dpeluffo@yachaytech.edu.ec

A. C. Umaquinga-Criollo
e-mail: acumaquinga@utn.edu.ec

L. L. Lorente-Leyva
e-mail: leandro.lorente@sdas-group.com

D. H. Peluffo-Ordóñez
Yachay Tech University, Urcuquí, Ecuador

Corporación Universitaria Autónoma de Nariño, Pasto, Colombia

A. C. Umaquinga-Criollo
Universidad Técnica del Norte, Ibarra, Ecuador

E. N. Cabrera-Alvarez
Universidad de Cienfuegos, Cienfuegos, Cuba
e-mail: elita@ucf.edu.cu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_23

1 Introduction

Instance selection methods represent an important approach in different areas of data science. In [1], some important elements are considered in topics related to instance selection. There are two important processes, namely training set selection and prototype selection.

The selection of a subset of data for another very large data set is summarized in the concept of “data condensation” [2]. This form of data reduction differs from the others and is integrated as one of the families of the instance selection methods. Mainly, data condensation approaches are studied based on the classification processes, particularly the k-nearest neighbor (KNN) methods which refer to obtain a consistent minimum set that classifies the entire original set. Figure 1 shows a simple representation of the KNN.

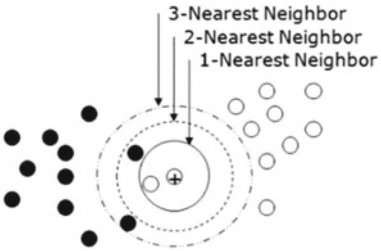
One of the first pioneering methods in the analysis in the data structure for the selection of instances was CNN [4]. The methods of condensation of data that are not related to the classification process are also known as methods of condensation of generic data, such condensation is performed through the so-called vector quantization (VQ), and example of this is the self-organization map and other ways of organizing the data as shown in Fig. 2.

1.1 Vector Quantization

Vector quantization (VQ) is a classic method that consists of approximating a continuous probability density function $p\ (x)$ of the vector input variable x by using a finite number of book-encoded vectors mi, $i = 1, 2, \ldots, k$; once these book-encoders have been chosen, the approximation of x implies finding the reference vector closest to x. An optimal location type of m minimizes to E where E is the rth power of the reconstruction error [5]:

$$E = \int \|x - m_c\|^r p(x)\mathrm{d}x \tag{1}$$

Fig. 1 K-nearest neighbor representation [3]



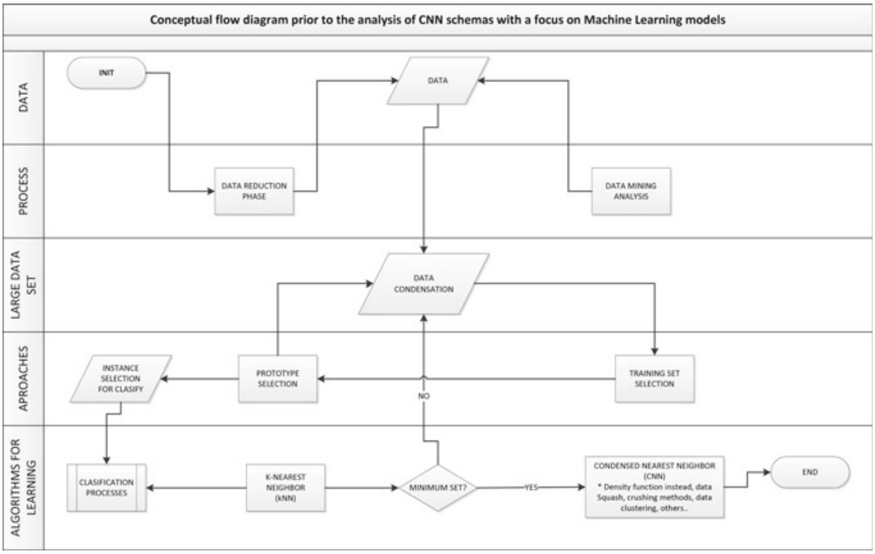

Conceptual flow diagram prior to the analysis of CNN schemas with a focus on Machine Learning models
DATA
PROCESS
LARGE DATA SET
APROACHES
ALGORITHMS FOR LEARNING
INIT
DATA
DATA REDUCTION PHASE
DATA MINING ANALYSIS
DATA CONDENSATION
INSTANCE SELECTION FOR CLASIFY
PROTOTYPE SELECTION
TRAINING SET SELECTION
NO
CLASIFICATION PROCESSES
K-NEAREST NEIGHBOR (kNN)
MINIMUM SET?
YES
CONDENSED NEAREST NEIGHBOR (CNN)
* Density function instead, data Squash, crushing methods, data clustering, others..
END

Fig. 2 CNN decision diagram for data reduction task

where dx represents the differential volume in the space x and the index $c = c\,(x)$ of the best match between the book-encoders (winner) is a function of the input vector:

$$\|x - m_c\| = \underbrace{\min}_{i}\{\|x - m_i\|\} \tag{2}$$

In general, a closed solution for the optimal location of m is not possible, so iterative approximation schemes can be used.

1.2 Condensed Methods

Generic data condensation methods are based on techniques that consider density; they consider the density function instead of minimizing the quantification error; that is, for a specific input set, the condensed output set [6] is established.

Other methods such as data squash or data clustering are used for sample selection. A crushing method seeks the compression of the data in such a way that a statistical analysis performed on the compressed data obtains the same result as with the original data. Clustering-based algorithms [7, 8] divide data into samples like each other and different from examples of data belonging to other groups [1].

Figure 3 is represented according to a distance function where the quality of the cluster could be measured according to the dimension of its diameter which is the maximum between two samples belonging to the same group.

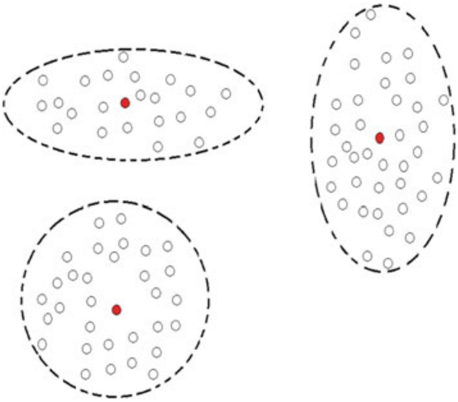


Fig. 3 Three clusters obtained from a set of two-dimensional data

1.3 Machine Learning and Feature Selection

In machine learning, a process known as feature selection consists of the selection of characteristics, attributes or selection of variable subsets for use in model building. In [2], two feature selection strategies are mentioned, the first based on feature ranking and the other based on best subset selection. In the case of the methods based on feature ranking, some statistical metrics are used, some of the simple complexity uses the correlation coefficient instead of other more complex used methods such as the Gini index, and this index can be used to quantify inequalities in variable distributions. Other feature ranking methods mentioned in the literature [9] are the bivariate and multivariate methods; these methods calculate the distance between the actual joint distributions of the characteristics of two or more variables and answer the question of what the joint distribution would be if these variables were independent, further. The joint distribution represents the probability distribution of existing case studies. Among the multivariate analysis, methods are the stepwise linear regression [10, 11] which has been used in cluster tasks and sample selection [12]; other slightly more complex algorithms include the use of machine learning and advanced statistics, for example, partial least squares regression [13] and sensitivity analysis [14]. Also, in performance analysis of virtual clusters [15] and architecture in wireless networks [16].

The second strategy based on subset selection has its focus on the selection of a subset for the selection of characteristics or attributes that have a significant effect on the prediction of a variable. The classic methods of data reduction and sample selection [17] mention its importance given the analysis of large amounts of data for each sample and the time consumed which may cause an over-adjustment of the model of training.

In all the approaches seen so far in a very simple way, the importance of selecting a suitable sample has been evidenced to reduce computational cost and time among other aspects. From now on, the various efforts made to obtain results using the

CNN method [18] with the prototype approach that facilitates the machine learning approach [19] will be more rigorously required.

The rest of this paper is structured as follows: Sect. 2 presents the theoretical background and overview referring to the main problem by the CNN method. Section 3 describes more practically by introducing the idea of the use of metrics in unsupervised learning and its relationship with CNN. Finally, the conclusions are presented in Sect. 4.

2 Theoretical Background and Overview

In practical problems, one of the most important elements to handle is the elimination of noise, redundancies, useless instances and therefore the selection of prototypes, constituting the first step for any practical application.

2.1 Problem Definition

It is desired to isolate the smallest set of instances that could predict the class with the same or greater precision than the original set [20]:

Lemma 2.1.1. *Let* X_p*be an instance where* $X_p = \left(X_{p1}, X_{p2}, ..., X_{pm}, X_{pc}\right)$*, with* $X_p \in c$*given by* X_{pc}*and a* $X_{pi} \in R^m$*being the value of the ith feature of the* p_{th}*sample. A training set TR, and also the N instances* X_p*and a validation set TS with t instances* X_p*, is obtained.* $S \subset TR$ *is the subset of the selected samples that resulted from applying an instance selection algorithm.*

Summarizing Lemma 2.1.1.*, the objective of an instance selection method is to obtain a subset* $S \subset T$ *such that S does not contain unnecessary instances* [21]*:*

$$\mathrm{Acc}(S) \cong \mathrm{Acc}(X) \tag{3}$$

where Acc(X) is the qualifier of the training set X.

2.2 Prototype-Based Approach on Unsupervised Learning

Models based on prototype analysis represent several appealing concepts such as the explicit representation of observations, data or typical representatives that exhibit some relation to psychology and neuroscience.

In Sect. 1, the relationship between condensation methods and vector quantization was approached in a very simple way, and this subsection discusses how to prototype selection matches the instance selection approach with a competitive perspective in unsupervised learning [18].

The vector quantization mathematical statement is formulated in terms of a function that represents costs and generally guides the computation of prototype vectors. A prototype-based representation [22] of a given set of P is defined in Lemma 2.2.1.

Lemma 2.2.1. *Assign the representation of a set of P feature vectors* $\{x^\mu \in \mathbb{R}^n\},\ \mu = 1, 2, ..., P$ *that represent a particular input values.*

A popular approach considers the assignment of any data point to the closest prototype, the so-called winner in the set $W = \{w^1, w^2, \ldots, w^K\}$ *in terms of a predefined distance measure.*

Using the Euclidean metric in feature space with:

$$d^2(x, y) = (x - y)^2 \text{ for } x, y \in \mathbb{R}^N \tag{4}$$

Having the quantization error [3] as the corresponding cost function:

$$H_{\mathrm{VQ}} = \sum_{i=1}^{P} \frac{1}{2} d^2\big(w^*(x^\mu), x^\mu\big) \tag{5}$$

where $w^*(x^\mu) \in \mathbb{R}^N$ denote the closest prototype using a Euclidean metric $x^\mu \in \mathbb{R}^n$:

$$d\big(w^*(x^\mu), x^\mu\big) \le d\big(w^j, x^\mu\big) \text{ for all } j = 1, 2, \ldots, K \tag{6}$$

The quantization error quantifies the fidelity with which the set of prototypes represent data.

2.3 The Condensed Nearest Neighbor Rule (CNN Rule)

An in-depth study on the pillars that support the CNN method [23] and that will be specified below:

Let $\big(X_1^{'}, Y_1^{'}\big) \ldots \big(X_m^{'}, Y_m^{'}\big)$ be a sequence that depends somehow on the data D_n, and let g_n be the 1-nearest neighbor rule with $\big(X_1^{'}, Y_1^{'}\big) \ldots \big(X_m^{'}, Y_m^{'}\big)$ where m is previously set. One way to find the data is to find the subset of the size m data, for the remained minimal $n - m$ data is confirmed by the error with the I-NN rule (this is known as Hart's rule).

If:

$$\hat{L}_n = \left(\frac{1}{n}\right) \sum_{i=1}^{n} I_{\{g_n(X_i) \neq Y_i\}} \tag{7}$$

And:

$$L_n = P\{g_n(X) \neq Y|D_n\} \tag{8}$$

Then, we have the following:

Lemma 2.3.1. *∀ $\varepsilon > 0$,*

$$P\left\{|L_n - \hat{L}_n| \geq \varepsilon\right\} \leq 8\mathrm{e}^{-\frac{n\varepsilon^2}{32}}\left(\frac{ne}{d+1}\right)^{(d+1)m(m-1)} \tag{9}$$

where $\hat{L}_n$is about the estimate error probability.

Observe that:

$$\hat{L}_n = \left(\frac{1}{n}\right)\sum_{i=1}^{n} I_{\left\{(X_j,Y_j)\notin\bigcup\limits_{i=1}^{m} B_i\times\left\{Y_i'\right\}\right\}} \tag{10}$$

where B_i is the Voronoi cell of $X_i^{'}$ corresponding to $X_1^{'} \ldots X_m^{'}$, where $B_i \subset R^d$ is the closer partition to $X_i^{'}$ than to any other $X_j^{'}$:

$$L_n = P\left\{(X, Y) \notin \bigcup_{i=1}^{m} B_i \times \left\{Y_i^{'}\right\}|D_n\right\} \tag{11}$$

Using simple upper bound:

$$|L_n - \hat{L}_n| \leq \underbrace{\mathrm{Sup}}_{A\in A_m}|v_n(A) - v(A)| \tag{12}$$

where v denotes the measure of (X, Y), v_n is some measure and A_m refer a set of all subsets of $R^d \times \{0, 1\}$ of the form $\bigcup_{i=1}^{m} B_i \times \{y_i\}$ where $B_1, \ldots, B_m$ are Voronoi's cells corresponding to $x_1, \ldots, x_m, x_i \in R^d, y_i \in \{0, 1\}$.

Using the Vapnik–Chervonenkis inequality [24]:

$$s(A_m, n) \leq s(A, n)^m \tag{13}$$

Such that A is the class of sets $B_1 \times \{y_1\}$ and each set in A intercepts in at most $m - 1$ hyperplanes. Then:

$$s(A, n) \leq \underbrace{Sup}_{n_0,n_1:n_0+n_1=n}\left(\prod_{j=0}^{1}\left(\frac{n_j e}{d+1}\right)\right)^{(d+1)(k-1)} \leq \left(\frac{n_j e}{d+1}\right)^{(d+1)(k-1)} \tag{14}$$

where n_j denotes the points $R^d \times \{j\}$ and the result follows from the Vapnik–Chervonenkis.

Other condensate rules based on CNN were also presented in [25, 26].

Table 1 New approaches based on traditional CNN methods

Method	Short description	References
Extended nearest neighbor	Used for pattern recognition	[10]
The fast-condensed nearest neighbor algorithm	Reuse Voronoi's concepts	[18]
Hierarchy extreme learning machine, for instance, selection	Fuzzy c-means utilizes condensed nearest neighbor (CNN) to make a preliminary selection of training samples	[7]
A modified firefly algorithm for image classification	Used in image classification task	[8]
Nonparametrically regression algorithm with instance selection	Provide flexible forms of prediction	[11]

An approach to the CNN algorithm [27, 28] can be as follows:

```
Algorithm 1
$T \leftarrow \emptyset$
Do
$\forall x \in X(inrandomorder)$
Find $x' \in T$ such that
$x - x' = \underbrace{min}_{x^w \in T} x - x^w$
If $Class(x) \neq Class(x^w)$ insert x to T
While T does not change
```

Several investigations have been carried out to interpret, extend and enhance the traditional CNN algorithm [29, 30]. In Table 1, some novel variants of implementation and application of the CNN method are shown.

3 Results and Discussion

A small review of the process of selecting instances has shown the high potential of sample selection techniques. Its application is valid in all areas and sub-areas of the modern world. The prototyping approach given by machine learning contributes too many investigations to reduce the computational cost of processes and the tasks of classifying huge amounts of data. Stopping in the analysis of the condensed nearest neighbor (CNN) algorithm [31], it represents a cognitive and theoretical element that means the basis of other evolutionary models.

The CNN algorithms use one nearest neighbor rule to iteratively decide if a sample should be removed or not [4].

3.1 Metric Considerations and Visual Scheme for the CNN

Many unsupervised algorithms perform unsupervised learning of distance metrics using information from the data itself or from the dimension where they are represented. In the selection of instances, the measurement of the distance between instances or the metric used is of crucial importance.

To formalize, denote the vectors x and y to those that represent the attributes of two instances x and y (classes are excluded).

A widely used metric is the Minkowski metric, which is defined as:

$$d = \sqrt[p]{\sum_{j=1}^{m} d_j^p} \tag{15}$$

where d_j is defined for continuous attributes such as $d_j = \left|x_j - y_j\right|$.

For some values of p, the Minkowski distance corresponds to a special metric as reflected in Table 2.

There are other important metrics such as the Mahalanobis distance based on the location of multivariate outliers to indicate an unusual combination between one or more variables.

A simple definition to this problem [10] is defined by:

$$d(\text{Mahalanobis}) = \left[(x_B - x_A)^{\mathrm{T}} * C^{-1} * (x_B - x_A)\right]^{0.5} \tag{19}$$

where:

x_A and x_B are a pair of objects and C is the sample covariance matrix.

The following figure shows some examples of sample selection using the Euclidean and the Mahalanobis distance using the CNN algorithm and comparing some values for the n-neighbors:

Figure 4 shows the importance of the selection and use of metrics at the time of clustering, as indicated by the classic methods of selection of instances, and the

Table 2 Minkowski metrics for different p values

Minkowski variant metric	The p value	Metric
Manhattan distance	1	$d = \sum_{j=1}^{m} d_j^p$ (16)
Euclidean distance	2	$d = \sqrt{\sum_{j=1}^{m} d_j^p}$ (17)
Chebyshev distance	∞	$d = \underbrace{\overbrace{\text{Max}}^{n}}_{j=1} \left|d_j\right|$ (18)

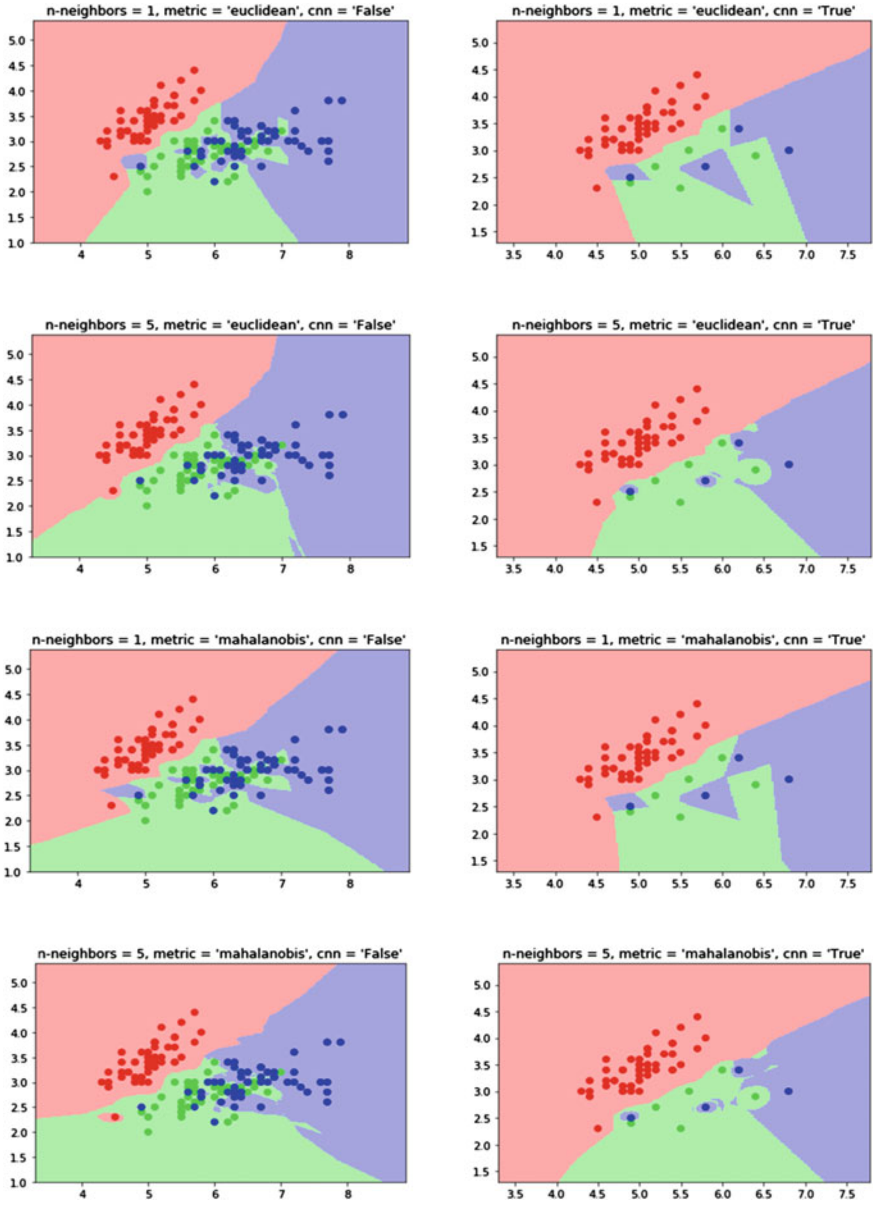


Fig. 4 Sample selection considering the Euclidean and Mahalanobis distance

fact of resorting to a sample that is sufficiently representative of a large population constitutes a difficult job. In this case, the example presented in Fig. 4 shows how the red, green and blue points are selected reflecting their color in a determined area according to the Euclidean and Mahalanobis metrics but using the CNN algorithm (squares on the right in Fig. 4) or simply using the aforementioned metrics (left squares in Fig. 4). As can be seen, using the CNN algorithm in combination with one of the two metrics achieves a clearer and more precise level of the reduced instances.

4 Conclusion

The beginning of the history of instance selection algorithms can be placed in the CNN algorithm (condensed nearest neighbor rule) whose contribution is due to Hart in 1968. The algorithm in its simplest state leaves in S a subset of T such that each element of T is closer to an element of S of the same class than to an element of S of a different class. From this idea, various variants have been formulated with an elegant mathematical profile that has allowed the reduction of computational costs in various modern problems given its simplicity.

Finally, the aim of this work has been to show some theoretical elements about the importance of the sample selection process and the condensed nearest neighbor method collected in the effort of several authors who have tried to theorize in complex aspects of the real world to give solutions to problems of today's world.

Acknowledgements The authors are greatly grateful for the support given by the SDAS Research Group. https://sdas-group.com/.

References

1. García S, Luengo J, Herrera F (2015) Data preprocessing in data mining. Springer International Publishing, Cham
2. Nisbet R, Elder J, Miner G (2009) Handbook of statistical analysis and data mining applications. Elsevier
3. Liu B (2011) Web data mining: exploring hyperlinks, contents, and usage data, 2nd edn. Springer, Heidelberg, New York
4. Hart P (1968) The condensed nearest neighbor rule (Corresp.). IEEE Trans Inf Theory 14:515–516. https://doi.org/10.1109/TIT.1968.1054155
5. Kohonen T (1990) The self-organizing map. Proc IEEE 78:1464–1480. https://doi.org/10.1109/5.58325
6. Girolami M, He C (2003) Probability density estimation from optimally condensed data samples. IEEE Trans Pattern Anal Mach Intell 25:1253–1264
7. Tang B, He H, Zhang S (2020) MCENN: a variant of extended nearest neighbor method for pattern recognition. Pattern Recogn Lett S0167865520300143. https://doi.org/10.1016/j.patrec.2020.01.015
8. Dey N (2020) Applications of firefly algorithm and its variants: case studies and new developments. Springer, Singapore

9. Chen Y, Liu Y, Ning J, Nie L, Zhu H, Chu H (2017) A composite likelihood method for bivariate meta-analysis in diagnostic systematic reviews. Stat Methods Med Res 26:914–930. https://doi.org/10.1177/0962280214562146
10. Stephanie (2017) Mahalanobis distance: simple definition, examples. In: statistics how to. https://www.statisticshowto.com/mahalanobis-distance/. Accessed 19 July 2020
11. Gong C, Wang P, Su Z (2020) An interactive nonparametric evidential regression algorithm with instance selection. Soft Comput. https://doi.org/10.1007/s00500-020-04667-4
12. Silhavy P, Silhavy R, Prokopova Z (2017) Evaluation of data clustering for stepwise linear regression on use case points estimation. Adv Intell Syst Comput 575:491–496. https://doi.org/10.1007/978-3-319-57141-6_52
13. Biancolillo A, Næs T (2019) The sequential and orthogonalized PLS regression for multiblock regression. In: Data handling in science and technology. Elsevier, pp 157–177
14. Barraza N, Moro S, Ferreyra M, de la Peña A (2019) Mutual information and sensitivity analysis for feature selection in customer targeting: a comparative study. J Inf Sci 45:53–67. https://doi.org/10.1177/0165551518770967
15. Smys S, Bala GJ (2012) Performance analysis of virtual clusters in personal communication networks. Cluster Comput 15:211–222. https://doi.org/10.1007/s10586-012-0209-8
16. Jyothirmai P, Raj J, Smys S (2017) Secured self organizing network architecture in wireless personal networks. Wireless Pers Commun 96:5603–5620. https://doi.org/10.1007/s11277-017-4436-4
17. Xu X, Li S, Liang T, Sun T (2020) Sample selection-based hierarchical extreme learning machine. Neurocomputing 377:95–102. https://doi.org/10.1016/j.neucom.2019.10.013
18. Ros F, Guillaume S (2020) Sampling techniques for supervised or unsupervised tasks. Springer International Publishing, Cham
19. Cerruela-García G, de Haro-García A, Toledano JP-P, García-Pedrajas N (2019) Improving the combination of results in the ensembles of prototype selectors. Neural Netw 118:175–191. https://doi.org/10.1016/j.neunet.2019.06.013
20. Brighton H, Mellish C (2002) Advances in instance selection for instance-based learning algorithms. Data Min Knowl Disc 6:153–172. https://doi.org/10.1023/A:1014043630878
21. Garcia S, Derrac J, Cano JR, Herrera F (2012) Prototype selection for nearest neighbor classification: taxonomy and empirical study. IEEE Trans Pattern Anal Mach Intell 34:417–435. https://doi.org/10.1109/TPAMI.2011.142
22. Biehl M, Hammer B, Villmann T (2016) Prototype-based models in machine learning: prototype-based models in machine learning. WIREs Cogn Sci 7:92–111. https://doi.org/10.1002/wcs.1378
23. Devroye L, Györfi L, Lugosi G (1996) A probabilistic theory of pattern recognition. Springer, New York, NY
24. Blumer A, Ehrenfeucht A, Haussler D, Warmuth MK (1989) Learnability and the Vapnik-Chervonenkis dimension. J ACM 36:929–965. https://doi.org/10.1145/76359.76371
25. Gates W (1972) The reduced nearest neighbor rule
26. Fukunaga K, Mantock JM (1984) Nonparametric Data Reduction. IEEE Trans Pattern Anal Mach Intell PAMI-6:115–118. https://doi.org/10.1109/TPAMI.1984.4767485
27. Ullmann J (1974) Automatic selection of reference data for use in a nearest-neighbor method of pattern classification (Corresp.). IEEE Trans Inform Theory 20:541–543. https://doi.org/10.1109/TIT.1974.1055252
28. Ritter G, Woodruff H, Lowry S, Isenhour T (1975) An algorithm for a selective nearest neighbor decision rule (Corresp.). IEEE Trans Inform Theory 21:665–669. https://doi.org/10.1109/TIT.1975.1055464
29. TOMEK I (1976) Two modifications of CNN. IEEE Trans Syst, Man, Cybern SMC-6:769–772. https://doi.org/10.1109/TSMC.1976.4309452
30. Swonger CW (1972) Sample set condensation for a condensed nearest neighbor decision rule for pattern recognition 511–519
31. Gowda K, Krishna G (1979) The condensed nearest neighbor rule using the concept of mutual nearest neighborhood (Corresp.). IEEE Trans Inform Theory 25:488–490. https://doi.org/10.1109/TIT.1979.1056066

Virtual Group Movie Recommendation System Using Social Network Information

Tranos Zuva and Keneilwe Zuva

Abstract Recommendation systems (RS) are software tools and methods designed to give recommendations to support customers in different decisions in terms of what items to buy, music to listen to, news to read, and so forth. Most recommender systems recommend items in terms of individual user likings and group recommender systems recommend items taking into consideration the likings and personalities of group members. To generate effective recommendations for a group, the system must satisfy, to the greatest extent possible, the individual interests of the group members. With the social networks, it is possible to recommend to a virtual group thus this study endeavors to develop a virtual group recommender system prototype using a model-based matrix factorization algorithm of collaborative filtering technique then popularity vote for virtual group. A publicly available dataset was used in this study. The results of the prototype showed the proposed collaborative filtering algorithm for prediction of user rating preferences demonstrated a good mean average error (MAE) of 0.70 and root mean square error (RMSE) of 0.89. Virtual groups of social networks user were then formed using the popularity vote algorithm and the results were plausible. This type of recommendation to a virtual group also enables members of the group to have something to talk about on the social network.

Keywords Mean absolute error (MAE) · Root mean squared error (RMSE) · Recommendation system (RS) · Collaborative-based filtering (CF) · Content-based filtering (CBF)

1 Introduction

With the rapid development of the Internet, more and more online services inevitably suffer from information overload, which makes it very hard for users to find the

---

T. Zuva (✉)
ICT Department, Vaal University of Technology, Vanderbijlpark 1900, South Africa
e-mail: tranosz@vut.ac.za

K. Zuva
Computer Science Department, University of Botswana, Gaborone, Botswana

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_24

information they need. Recommendation systems have proved to be effective means of dealing with the knowledge overload for online users and have become one of the most important and popular resources in electronic commerce [1]. Recommendation systems help users to identify items that match a user's needs and preferences from a generally long list of possibly interesting items. In literature, several recommendation techniques have been proposed [1].

Content-based filtering (CBF) endeavors to recommend almost identical items to the ones the consumer has preferred before; this is done through considering the items' features [2]. The usual approach is to use the same feature space to represent both the consumers and the items. The similarity scores are then computed between the user's profile and items' profiles. The similarity scores are then used for the recommendation of items to the user concerned. The algorithm performs perfectly well for users that do not have a lot of historical data that can be used during the recommendation period.

Collaborative-based filtering (CF) algorithm is an algorithm that is used for prediction of ratings of an item of interest of a user using many related users of similar tastes (collaborating). Collaborative filtering recommendation systems use a database about user preferences to predict additional items a new user might have interest in [3]. Collaborative filtering, seen as one of the most popular algorithms in developing recommendation systems applications, predicts the unknown preferences ratings using the established preferences ratings of a group of consumers [1].

Hybrid recommendation systems as the name entails are based on the combination of recommender systems algorithms such as content-based and collaborative-based filtering techniques. A hybrid system combines collaborative and content-based filtering techniques and tries to use the advantages of CBF to ameliorate the disadvantages of CF. For example, CF experiences problems with new items, i.e., they have problems in recommending items that have no ratings. Content-based techniques can recommend new items, this is so because they predict using the descriptions (features) of the items that are usually available. The creation of new hybrid recommender systems can be done by combining two or more basic RS techniques in several ways [1, 4].

Social network sites are Web-based services that permit individuals to:

(a) Build a public or semi-public profile within a bounded system,
(b) Formulate a list of other users with whom they share a connection, and
(c) Display and browse their list of ties and those created inside the system by others.

The nature and classification of these connections can vary from one site to another [1]. With very little time and effort, user registered with a social networking site creates his profile which contains some basic details. A social network user can do things such as adding new friends, uploading images and/or audio/videos, setting status messages, making comments, joining various groups of people who share similar interests, and joining forums for discussion but not limited only to the ones listed.

Although most recommendation systems recommend items according to an individual consumer's preferences, group recommendation systems propose items that take into account the group members' preferences and personalities [5]. To produce appropriate recommendations for a group, the program must meet the individual needs of the members of the group as much as possible [1].

2 Background

A variety of recommendation generation techniques, including content-based filtering (CBF), collaborative filtering (CB), and hybrid recommendation systems, have been proposed. These techniques assist users in finding items of their choices such as services, products, or information. Through aggregating and reviewing recommendations from other users, feedback from different authorities, and user characteristics, most of these recommender systems on the market suggest digital goods, books, Web sites, music, movies, and TV shows, to name only a few. Collaborative filtering (CF) makes suggestions to consumers according to other consumers' ratings on items, putting more weights on those from similar consumers [6]. It is considered as one of the most successful recommendation techniques.

Most of the recommendation systems on the market recommend personal items to individuals rather than to a group of people to participate or use in a group [7]. In some recommendation domains, it is necessary and suitable to recommend items to individuals such as shopping and asset investment. In these domains, personal interests and behavior are very important in personal recommender systems. In other domains where a group of individuals require an item such as a movie(s), trip(s), book club(s), and restaurant(s) for use as a group require aggregating individual consumers' likings into a group's preference properly. Choosing the aggregation algorithm is the most daunting task in group recommender systems.

From this information, it is clear that content-based filtering has some limitations. These limitations can be countered by using the strengths of collaborative filtering to solve the weaknesses of content-based filtering. The study investigates existing filtering techniques, hence applying the collaborative filtering technique to recommend movies for a group of participants.

This research, therefore, attempts to make use of the advantages of the collaborative filtering technique to overcome the limitations of the content-based filtering technique to recommend movies to individuals, and eventually making movie suggestions to diversified virtual groups using information gathered from the social network, Facebook.

Recommendation systems are available in many Web applications to help the user in making their choices. They improve sales and are of benefit to businesses, but whether they benefit customers/users by providing relevant products is still questionable [8]. The relevance of this proposal is to assist diversified groups of users by engaging them in the process of movie selection, acquiring information from their social network profiles, making individual movie recommendations, and ultimately making movie suggestions for their respective diversified groups to meet an acceptable level of member satisfaction.

Movie recommendation systems have been beneficial to viewers for years. The virtual group movie recommendation system will be of paramount importance and change user's perception by providing the necessary information, not just to a single user, but also to diversified groups of viewers who will be watching movies together, in form of virtual groups. The system will carry out a series of calculations to reach an acceptable level of satisfaction for all the virtual group members. The objective is that with time, and as technology evolves, the proposed virtual group movie recommendation system for will be updated by either adding or removing features so that it may adapt to technological changes.

The utilization of recommendation systems, specifically movie recommendations systems, has been studied with various recommendation systems approaches employed. However, this study is expected to make a major contribution to how recommendation systems recommend movies to users. The study further focuses on how individual user ratings can be considered when making group movie recommendations. In addition, the study attempts to form virtual groups based on the similarities of users in terms of their ratings. The aim is that with these formed virtual groups, a movie will be recommended to each group.

3 Methodology

In this study, a model-based matrix factorization algorithm of collaborative filtering technique was used. In this perspective, the algorithm was deployed because of its accuracy in making predictions and because of its ability to improve prediction performance [9]. In addition, this algorithm has been proved to be a better option to address the issues of data sparsity, over-fitting, and convergence speed [10].

The general approach of the virtual group movie recommendation system followed these steps:

1. Prediction of movie ratings using matrix factorization
2. The standard ranking of movies above a pre-set threshold value (3.0)
3. Recommendations of three movies to individuals
4. Plurality check
5. Generation of virtual groups
6. Standard ranking to generate a group recommendation list
7. Recommendation of movies to a virtual group, together with a list of group members.

The approach followed by the virtual group movie recommendation system is illustrated in Fig. 1.

As illustrated in Fig. 1, the recommendation system predicts ratings for all unrated movies using matrix factorization algorithm of collaborative filtering technique. The

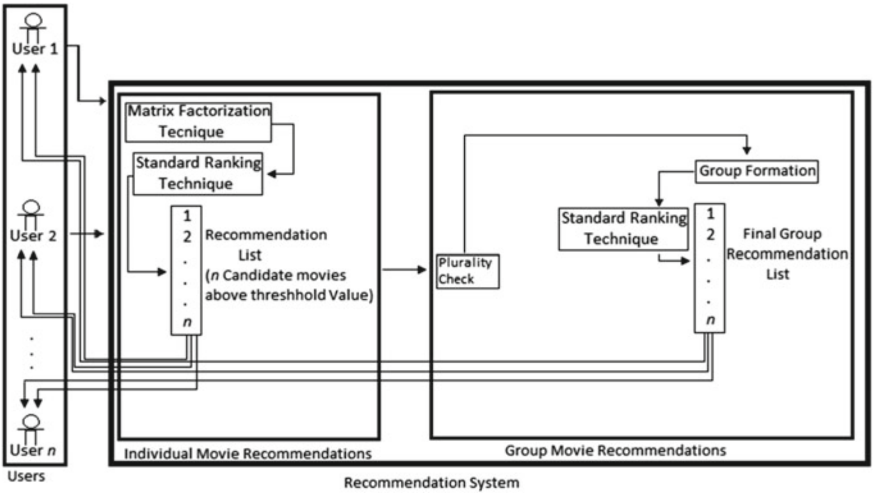

User 1
User 2
User *n*
Users
Matrix Factorization Tecnique
Standard Ranking Technique
1
2
.
.
.
n
Recommendation List (*n* Candidate movies above threshhold Value)
Plurality Check
Group Formation
Standard Ranking Technique
Final Group Recommendation List
Individual Movie Recommendations
Group Movie Recommendations
Recommendation System

Fig. 1 Virtual group movie recommendation system architecture derived from [1]

system then uses a standard ranking technique to determine top n predicted movies, after which the top n predicted movies that are above the pre-set threshold are sent to each user. Based on common movies recommendations with a predicted rating above the pre-set threshold, the systems go through the plurality check, generates virtual groups, undergoes standard ranking technique, recommends top n movies to the generated virtual groups, and sends movie recommendations to each group member. The list is sent together with the list of group members.

Prediction of movie ratings using matrix factorization that approximates matrix X by the product of two smaller matrices W and H, i.e., $X \approx WH^{\mathrm{T}}$. In terms of recommendation systems, the matrix X is the partially observed rating matrix, $W \in \mathbb{R}^{U \times K}$ is a matrix where each row i is a vector containing the K features describing the item i. Let $w_{uk} \in W$ and $h_{ik} \in H$, then the user u predicted rating of item i is calculated using Eq. 1.

$$\hat{r}_{ui} = \sum_{k=1}^{K} w_{uk} h_{ik} = \left(WH^{\mathrm{T}}\right)_{u,i} \tag{1}$$

where the model parameters are W and H, these can be learnt by optimizing a given criterion using stochastic gradient descent.

The standard ranking of movies above the pre-set threshold value is a commonly used approach for ranking the items in recommender systems. In this approach, the predicted rating of movies was ranked from highest to lowest. In our case, all the predicted movies were ranked in descending order above the pre-set threshold value, which was 3, which means that the highly predicted item comes first in the list and the lowest predicted item is at the bottom of the list. This is to guarantee that the suggested products are correct. This was achieved by utilizing Eq. 2 the same way it was used by [1].

$$\text{rank}_{\text{standard}}(i) = R * (u, i)^{-1} \tag{2}$$

where $R*(u, i)$ is the predicted rating. The power of -1 indicates that the items with the highest predicted are recommended to the user. This approach increases the accuracy in the recommendation system [11].

After movie rating predictions were computed using matrix factorization technique, and the standard ranking of movies to ensure accuracy was carried out, each user was sent a list of three recommended movies starting with the highest to the lowest predicted rating (descending). Figure 2 demonstrates an example of the steps carried out in generating the final movie recommendation list for each user.

Generation of virtual groups was the step. After the individual recommendations were made for each user, the recommendation system identifies four movies with the highest predicted rating (above threshold 3) to users. Four virtual groups were then formed out of these users based on the predictions above the threshold.

The plurality voting aggregation strategy was used in this study. The assumption is that the movie that is recommended to more people has the advantage of being watched by these people as a virtual group. The implication is that the virtual group movie recommendation system went through the plurality check to pick the four movies that have the highest number of votings (recommendations). All these movies were picked among the movies the users have not seen, and that have a prediction above the threshold. Table 1 shows an example of how the strategy was applied.

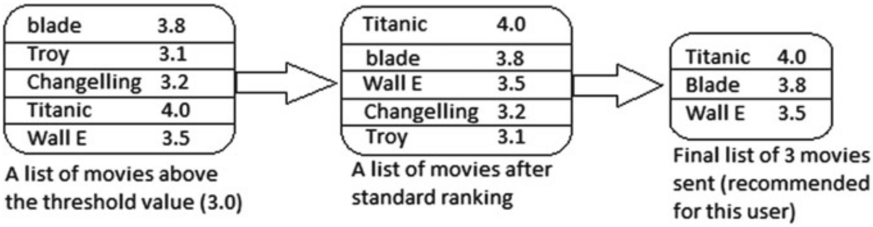

blade 3.8
Troy 3.1
Changelling 3.2
Titanic 4.0
Wall E 3.5
A list of movies above the threshold value (3.0)
Titanic 4.0
blade 3.8
Wall E 3.5
Changelling 3.2
Troy 3.1
A list of movies after standard ranking
Titanic 4.0
Blade 3.8
Wall E 3.5
Final list of 3 movies sent (recommended for this user)

Fig. 2 Individual movie recommendation list derived from[12]

Table 1 Plurality voting strategy example derived from [1]

Movie	User 1	User 2	User 3	User 4	User 5	User 6	Group average
Blade	4.0		4.8		5.0		4.6
Troy			4.5	3.3			3.9
Changeling	3.5	3.2		3.1	3.0		3.2
Titanic		4.5	3.5		3.2		3.7
Wall E				3.0		5.0	4.0

In the case of the example stated in Table 1, it is evident that the movie Blade has the highest group average of 4.6. But the plurality voting strategy would recommend the movie Changeling even though it has a group average of only 3.2. This would be done because the movie Changeling has the greatest number of votings (recommendations). It is worth noting that all the movies that go through this stage of plurality check are all predicted to be above the threshold.

On completion of the plurality check, the recommendation system undergoes the standard ranking process for group recommendations the same way it does with individual recommendations. Just like with individual recommendations list, this is done to increase accuracy. The list is displayed to virtual group members in descending order, starting with the movie with the highest group average to the lowest.

When both the plurality check and the standard ranking processes are complete, the recommendation list reaches its final stage where it sends the final list to each virtual group member as shown in Table 2. The final group movie recommendation list is sent to all the virtual group members together with the names/user ID's of all the group members as shown in Table 3.

Another essential part of this investigation was to evaluate the developed prototype to determine its effectiveness and accuracy in generating predictions and making recommendations. The evaluation metrics may be selected depending on the goal that the researcher wishes to achieve. To measure the predictive performance of the system, to obtain the error of the system during the implementation, the mean absolute error (MAE) and root mean squared error (RMSE) were calculated. These are the two most common metrics used to measure accuracy for continuous variables.

Mean absolute error (MAE): It was used to measures the average magnitude of the errors in a set of predictions, without considering their direction. As MAE measures accuracy for continuous variables, it was used to determine the average over the test

Table 2 Movies arranged in popularity

Movie	User 1	User 2	User 3	User 4	User 5	User 6	Group average	Popularity
Changeling	3.5	3.2		3.1	3.0		3.2	4
Blade	4.0		4.8		5.0		4.6	3
Titanic		4.5	3.5		3.2		3.7	3
Wall E				3.0		5.0	4.0	2
Troy			4.5	3.3			3.9	2

Table 3 Recommendations and virtual groups

Recommended movies	Virtual groups
Changeling	User 1, user 2, user 4, and user 5
Blade	User 1, user 3, and user 5
Titanic	User 2, user 3, and user 5
Wall E	User 4 and user 6
Troy	User 3 and user 4

sample of the absolute differences between prediction and actual observation where all individual differences have equal weight. Equation 3 illustrates how MAE was calculated.

$$\text{MAE} = \frac{1}{n}\sum_{j=1}^{n}\left|\mathcal{Y}_j - \hat{\mathcal{Y}}_j\right| \tag{3}$$

where $\hat{\mathcal{Y}}$ is the predicted rating, $\mathcal{Y}$ is the actual rating, and n is the number of occurrences/instances (amount of ratings).

The RMSE was also used to measure the accuracy of the prototype. Equation 4 was used for calculating the RMSE.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathcal{Y}_j - \hat{\mathcal{Y}}_j\right)^2} \tag{4}$$

where $\hat{\mathcal{Y}}$ is the predicted rating, $\mathcal{Y}$ is the actual rating, and n = number of occurrences/instances (the amount of ratings).

Both the MAE and RMSE can range from 0 to ∞ where ∞ is the maximum error depending on the rating scale of the measured application. They are negatively oriented scores, which mean lower values are better.

In this investigation, both the MAE and the RMSE were used together to diagnose the variation in the errors in a set of predictions. The RMSE always gives the larger or equal to the MAE; the greater the difference between them, the greater the variance in the individual errors in the sample. If the RMSE = MAE, then all the errors are of the same magnitude.

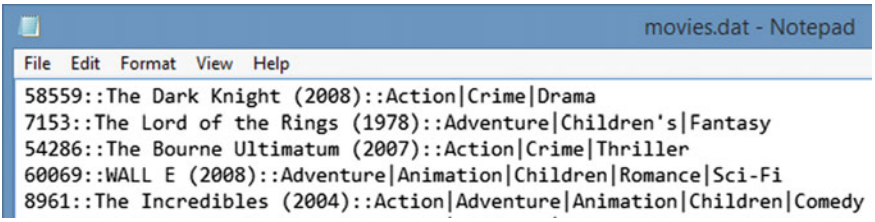
A publicly available dataset based on group recommender systems enhanced by social elements is constructed by Lara Quijano from the Group of Artificial Intelligence Applications (GIGA) obtained from (https://gaia.fdi.ucm.es/research/happymovie/download). The dataset consists of a sample of 58 users and 50 movies selected from the MovieLens dataset. Datasets were in form of two separate files, movies dataset and ratings dataset, in notepad files.

The movies file contains fifty movies that users had to rate movies he or she may have seen. This set of movies consists of a sample of all the different genres that existed and movie types so that with fifty movies a general idea of what types of movies a given user liked could be formed. The dataset had movie_id, movie_name, and genre attributes. Table 4 gives detailed descriptions of the attributes in the movies dataset while Fig. 3 shows an overview of the movies dataset Microsoft Excel file. Only the first ten elements of the movies dataset and its attributes are displayed.

The rating file had ratings ranging from 0.5 ratings as the minimum possible rating and 5.0 as the highest possible rating. The file had user_id, movie_id, and rating attributes. Table 5 gives detailed descriptions of the attributes in the rating dataset. Figure 4 shows an overview of the ratings dataset after the data. Only the first ten elements of the dataset are displayed here.

Table 4 Movies dataset attribute description table

Attribute	Description
MovieLens id number	A unique number identifying each movie in the dataset. Every single movie has a special movie_id for identification. movie_ids used here are similar to those used by grouplens in movielens
MovieLens name	This attribute holds the title of the movie. Similar to movie id, the movie names used in the dataset are identical to movie names used by grouplens in movielens
Year	This attribute holds a specific year a movie was released
Genre	A movie genre attribute refers to a motion-picture category based on similarities either in the narrative elements or in emotional response to the film. The genres contained in the dataset are Action, Adventure, Animation, Children, Comedy, Crime, Documentary, Drama, Fantasy, FilmNoir, Horror, Musical, Mystery, Romance, SciFi, Thriller, War, and Western



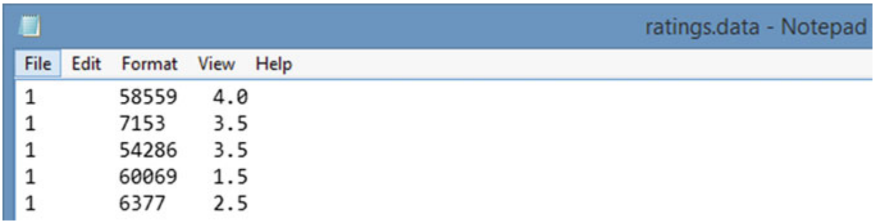
```
58559::The Dark Knight (2008)::Action|Crime|Drama
7153::The Lord of the Rings (1978)::Adventure|Children's|Fantasy
54286::The Bourne Ultimatum (2007)::Action|Crime|Thriller
60069::WALL E (2008)::Adventure|Animation|Children|Romance|Sci-Fi
8961::The Incredibles (2004)::Action|Adventure|Animation|Children|Comedy
```

Fig. 3 Overview of movies dataset notepad file

Table 5 Ratings dataset attribute description table

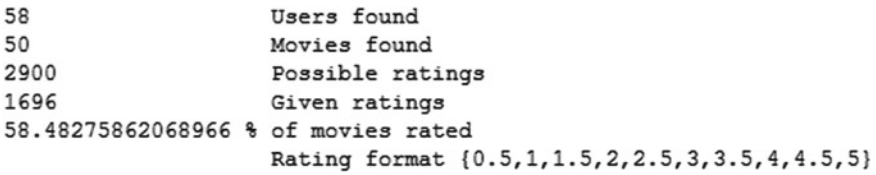
Attribute	Description
user_id	A unique number identifying the user in the dataset. Every single movie has a special user_id for identification. The dataset consists of 58 users. As a result, user_ids begin at 1–58
movie_id	A unique number identifying movies in the dataset. Every single movie has a special movie_id for identification. movie_ids used here are similar to those used by grouplens in movielens
rating	Rating attribute holds the ratings than 58 users gave to the 50 movies that presented them ratings range from 0.5 which is awarded for movies the user least liked to 5.0 for movies the user considered flawless in every department and left a long-lasting impression on them

When conducting this investigation, the author understood that data quality is important and without accurate, good-quality data, a significant amount of time, effort, and resources would be wasted trying to develop a recommendation system. As a result, it was ensured that only the most accurate and relevant data was entered and used in the datasets. Data scrubbing also referred to as data cleansing may be



```
1       58559   4.0
1       7153    3.5
1       54286   3.5
1       60069   1.5
1       6377    2.5
```

Fig. 4 Overview of ratings notepad dataset



```
58                   Users found
50                   Movies found
2900                 Possible ratings
1696                 Given ratings
58.48275862068966 %  of movies rated
                     Rating format {0.5,1,1.5,2,2.5,3,3.5,4,4.5,5}
```

Fig. 5 Results after preprocessing data

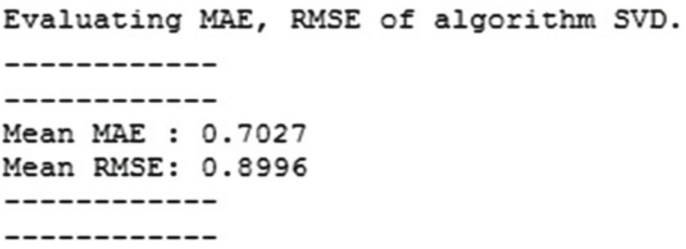
described as the identification of errors within a dataset, and the removal or correction of those errors. This process involves ensuring that your data is correct, consistent, and usable by identifying and removing or correcting any errors or corruptions in the data. After an establishment of which attributes to use, the irrelevant data was then left out and the movies and ratings datasets were loaded into the data frame, using python, in Jupiter notebook integrated development environment. To obtain meaningful data from the datasets, movies and ratings datasets were merged into one data frame using movie_id attribute.

After the data was cleaned, the author then preprocessed the data to obtain information from it as the raw data is non-comprehensive. By so doing, the author added a visual aspect to the data, making it easier and quicker to understand. This process included opening movies dataset (movies.dat) and ratings dataset (ratings.data), removing attributes that are not relevant to the study, removing or correcting errors from movies.dat and ratings.data and saving the changes on both movies.dat and ratings.data files.

After finalizing the preprocessing, the datasets contained 58 users and 50 movies with 1696 given ratings out of possible 2900 ratings. This implies that 58.5% ratings were given and the ratings expected were in a format 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 5.5, and 5.0. Figure 5 gives a general overview of the data after preprocessing.

4 Result Evaluation of the Prototype

Another crucial stage of the prototype was to evaluate its accuracy by comparing the predicted ratings directly with the actual ratings given by the users. To fulfill



```
Evaluating MAE, RMSE of algorithm SVD.
------------
------------
Mean MAE : 0.7027
Mean RMSE: 0.8996
------------
------------
```

Fig. 6 MAE and RMSE results of the prototype

this purpose, mean absolute error (MAE) and root mean squared error (RMSE) were deployed with the results shown in Fig. 6 [13]. A movie recommendation system using movielens dataset is proposed and 0.709531 MAE and 0.905520 RMSE using FunkSVD are achieved while achieving 0.717344 MAE and 0.9200979 RMSE using item-based collaborative filtering [14]. A movie recommendation system via K-means PSO-FCM technique is proposed that achieved 0.7547 as the MAE [1]. A system that achieved 0.82 MAE and 1.08 RMSE using fast maximum margin matrix factorization, 0.80 MAE and 1.05 RMSE using Iterative SVD, and 0.72 MAE and 0.95 RMSE using repeated matrix is proposed. Our prototype showed better performance over all these with 0.7027 MAE and 0.8996 RMSE as Fig. 6 illustrates.

The proposed collaborative filtering virtual group movie recommendation system using social network information has demonstrated a good 0.70 MAE and 0.89 RMSE. The algorithm has been explored and evaluated comprehensively. The findings depicted that the prototype fulfills its objectives and performs better than other recommendation systems considered for comparison.

5 Conclusion

In this article, a model-based matrix factorization algorithm of collaborative filtering technique was used to predict user preference scores for movies. The popularity vote ranking algorithm was then used for virtual group recommendation. The accuracy of the collaborative filtering technique was measured using mean absolute error and root mean squared error. The results were acceptable in comparison with other researchers thus gave us confidence in the results for the virtual group. The virtual group recommender system is believed to improve the interaction of users on social networks through discussion of the movies they would have watched at the same time. Our system might have to request users if they agree to let their contact details availed to people for those who have the same preference as theirs. The privacy issues will need to be addressed in future work.

References

1. Ricci F, Rokach I, Shapira B (2010)Introduction to recommender systems handbook.In: Nature S (ed) Recommender systems handbook. Bozen-Bolzano
2. Ricci F, Rokach L, Shapira B (2011) Introduction to recommender systems handbook. Springer, Boston
3. Ricci F, Rokach L, Shapira B (2015)Recommender systems handbook. In: Nature S (ed) Introduction and challenges. Springer International Publishing, Bolzano, pp 1–34
4. Lops P, Gemmis M, Semeraro G (2010)Recommender systems handbook. s.l.In: Nature S (ed) Content-based recommender systems: state of the art and trends. Springer International Publishing AG, pp 73–105
5. Breese J, Kadie CC (1998) Empirical analysis of predictive algorithms for collaborative filtering. Morgan Kaufmann Publishers Inc., Madison, Wisconsin
6. Zhongqi L (2015) Content-based collaborative filtering for news topic recommendation. AAAI Press ©2015, Austin, Texas
7. Ricci F, Rokach L, Shapira B (2010)Introduction to recommender systems handbook.In: Nature S (ed) Recommender systems handbook. Bozen-Bolzano, pp 1–35
8. Boyd DM, Ellison NB (2007) Social network sites: definition, history, and scholarship. J. Comput. Med. Commun. 1:210–230
9. Jameson A, Smyth B (2017)Recommendation to groups.In: Nature S (ed) The adaptive web. 2017 Springer International Publishing AG, Dublin, pp 596–627
10. Baltrunas L, Makcinskas T, Ricci F (2010) Group recommendations with rank aggregation and collaborative filtering. ACM, New York, NY, USA
11. Karunanithi N, Alspector J (1996)Feature-based and clique-based user models for movie selection. In: Proceedings of the fifth international conference, vol II, pp 130–136
12. Porcel C, Tejeda-Lorente A, Martínez MA, Herrera-Viedma E (2012) A hybrid recommender system for the selective dissemination of research resources in a Technology Transfer Office.Inf Sci 2:1–19
13. Al-Barznjl K, Atanassov A (2017) Collaborative filtering techniques for generating recommendations on big data. s.n., Sofia, Bulgaria
14. Kim JK, Kim KH, Oh YH, Ryu YU (2010) A group recommendation system for online communities. International Journal of information management. J Inf Prof 30:1–13

Hindi to English: Transformer-Based Neural Machine Translation

Kavit Gangar, Hardik Ruparel, and Shreyas Lele

Abstract Machine Translation (MT) is one of the most prominent tasks in Natural Language Processing (NLP) which involves the automatic conversion of texts from one natural language to another while preserving its meaning and fluency. Although the research in machine translation has been going on since multiple decades, the newer approach of integrating deep learning techniques in natural language processing has led to significant improvements in the translation quality. This paper has developed a Neural Machine Translation (NMT) system by training the Transformer model to translate texts from Indian Language Hindi to English. Hindi being a low resource language has made it difficult for neural networks to understand the language thereby leading to a slow growth in the development of neural machine translators. Thus, to address this gap, back-translation is implemented to augment the training data and for creating the vocabulary, it has been experimented with both word and subword level tokenization using Byte Pair Encoding (BPE) thereby ending up training the Transformer in 10 different configurations. This led us to achieve a state-of-the-art BLEU score of 24.53 on the test set of IIT Bombay English-Hindi Corpus in one of the configurations.

Keywords Neural machine translation · Transformer · Byte pair encoding · Back-translation

---

K. Gangar (✉) · H. Ruparel · S. Lele
Veermata Jijabai Technological Institute, Mumbai, India
e-mail: kavitgangar34@gmail.com

H. Ruparel
e-mail: hardikruparel14@gmail.com

S. Lele
e-mail: shreyaslele2398@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_25

1 Introduction

Machine translation is one of the oldest tasks taken up by computer scientists and the development in this field has been going on for more than 60 years. The research in this field has made remarkable progress to develop translator systems to convert source language to target language while maintaining the contextuality and fluency. In earlier times, the translation was handled by statically replacing words with the words from the target language. This dictionary look-up led technique led to inarticulate translation and hence was made obsolete by Rule-Based Machine Translation (RBMT) [1]. RBMT is a system based on linguistic information about the source and target languages derived from dictionaries and grammar including semantics and syntactic regularities of each language [2]. With the absence of flexibility and scalability to incorporate new words and semantics and the requirement of human expertise to define numerous rules, rule-based machine translation systems could only achieve accuracy on a subset of languages. To overcome the issues of the RBMT system, a new approach called Statistical Machine Translation (SMT) was introduced. Instead of having rules determine the target sequence, SMT approaches leverage probability and statistics to determine the output sequence. This approach made it feasible to cover all types of language within the source and target language and to add new pairs. Most of these systems are based on Bayesian prediction and have phrases and sentences as the basic units of translation. The main issue faced by this approach is the requirement of colossal amounts of data, which is a huge problem for low resource languages.

Due to these prevailing issues, there is a demand to explore alternate methods for creating a smarter and more efficient translation system. The development of various deep learning techniques and the promising results shown by the combination of these techniques with NLP created a new approach called NMT. NMT's advantage lies in two facts that are its simplistic architecture and its ability to capture long dependencies in the sentence, thereby indicating its huge potential in emerging as a new trend of the mainstream [3]. Conceptually speaking, NMT is a simple Deep Neural Network (DNN) that reads the entire source sentence and produces an output translation one word at a time. The reason why NMT systems are appealing is that they require minimal domain knowledge which makes it well-suited for any problem that can be formulated as mapping an input sequence to an output sequence. Also, the inherent nature of the neural networks to generalize any input implies that NMT systems will generalize to novel word phrases that are not present in the training set.

Moreover, almost all the languages in the world are continuously evolving with new words getting added, older words getting amended and new slangs getting introduced very frequently. Even though NMT systems generalizes the input data well, they still lack the ability to translate the rare words due to their fixed modest-size vocabulary which forces the NMT system to use *unk* symbol for representing out-of-vocabulary (OOV) words [4]. To tackle this issue, a subword tokenization technique called Byte Pair Encoding (BPE) was introduced. BPE divides the words such that the frequent sequence of letters is combined thereby forming a root word and affix. This approach alone handles the OOV words by merging the root word and the different combinations of affixes thereby creating the rare word [5].

This paper presents the experimental setup and the state-of-the-art results obtained after training the Transformer model on the IIT Bombay CFILT English-Hindi dataset of 1.5 million parallel records. The paper is organized as follows: Sect. 2 describes the motivation behind our work. Section 3, describes the implemented model. In Sect. 4 the details of the experimental setup is presented for training the model. In Sect. 5a comparative analysis of the results obtained by training the model is displayed in different configurations. Finally, Sect. 6 concludes the paper.

2 Motivation

With the power of deep learning, Neural Machine Translation has arisen as the most powerful approach to perform the translation task.

In [6] a model called Transformer was introduced which uses the encoder-decoder approach where the encoder first converts the original input sequence into its latent representation in the form of hidden state vectors. By using latent representation, can decode the predicted output sequence. Here transformed helps to achieve the parllelization to encode the symbol data position from the sequence.

In [7], explains about the improvement of machine translation for the monolingual series in the training data set. Dummy source sentence helps to detect the monolingual training instances and provides the better results in accuracy.

In [8], investigates the open vocabulary translation approach based on the byte pair encoding. This byte pair encoding is helps to compress the word segmentation and make it as open vocabulary with fixed sized by using neural network models.

Motivated with the results obtained by the transformer for machine translation on various languages, a translation system that translates a Hindi sentence to English is created. Since a low resource Hindi language is used for which the amount of good quality parallel corpus is limited, back-translation is applied to increase the quantity of training data. To overcome the problem caused by out of vocabulary words, BPE is used.

3 NMT Model Architecture

3.1 Structure

The Transformer model is the first NMT architecture that completely relies on the self-attention mechanism to calculate the representation of its input and output data without using recurrent neural networks (RNNs) or convolutional neural networks (CNNs) [9]. The Transformer model consists of an encoding unit and a decoding

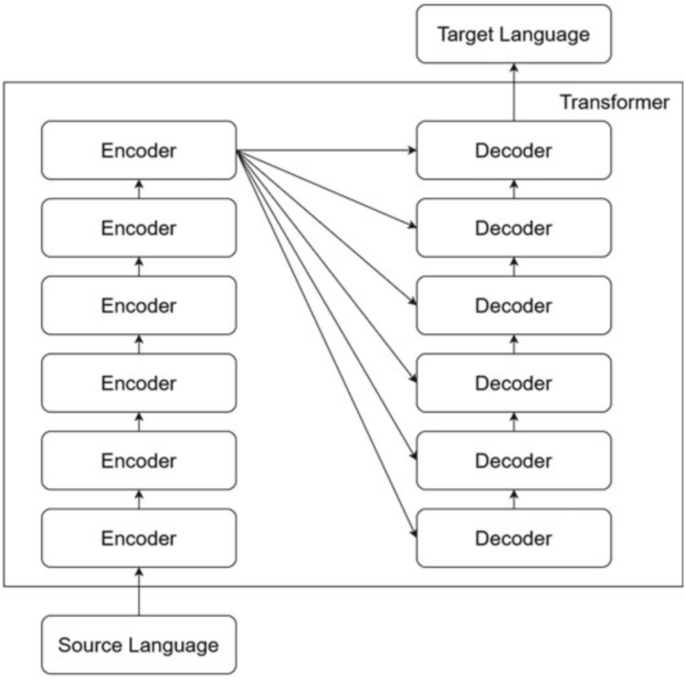


Fig. 1 Transformer structure—Bird's-eye view

unit wherein each of these components consists of a stack of 6 layers of encoders and decoders respectively. (see Fig. 1).

Each encoder layer consists of two sublayers. The first sublayer is the multi-head self-attention layer and the second sublayer is a position-wise fully connected feed forward network [6]. Each decoder layer in the decoding component consists of 3 sublayers. The function of the first two sublayers is the same as that in the encoder. However, the third sublayer performs multi-head attention mechanism over the output of the encoder stack (see Fig. 2).

3.2 Working

Before passing the input data to the encoder stack, the input data is first projected into an embedded vector space. To capture the notation, distance between different words and the order of the words in the input sequence, a positional embedding vector is added to the input vector. This intermediate vector is then fed to the first layer of the encoding component of the transformer. A multi-head self-attention is then computed on this intermediate embedded vector space. Multi-headed mechanism improves the performance of the attention-layer in two ways. First, it helps expand the model's

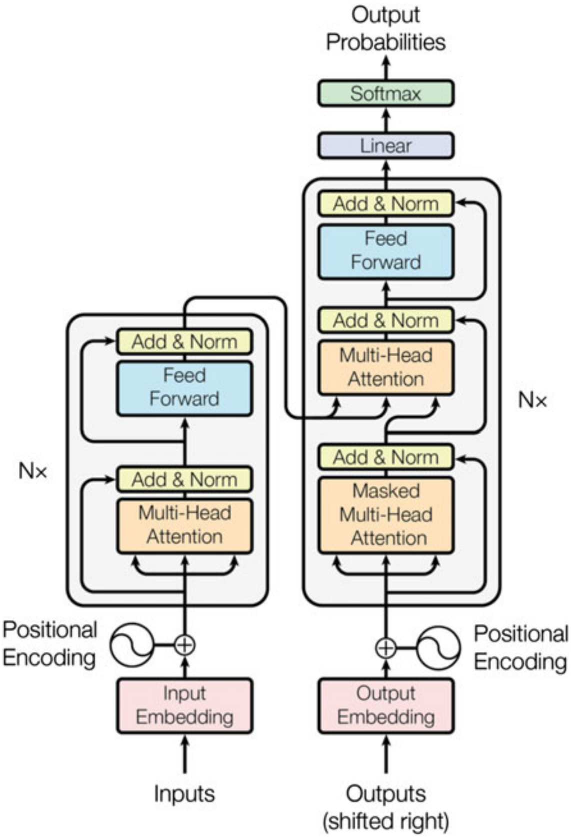


Fig. 2 The transformer architecture [6]

ability to focus on the words in different positions. Second, it gives the attention layer multiple representation subspaces, concatenates them and then projects linearly onto a space with initial dimensions [10]. The output of the self-attention layer is then passed onto a dense feed forward network which consists of two linear functions with RELU in between them. The output of this feed forward network is then passed on to another encoder layer stacked on top of it. All the encoder layers in the encoding component have the same functionality. Finally, the output of the encoding unit is then passed as an input to the decoding unit.

The decoder has similar functionality as that of the encoder. The output of the top encoder is converted into a set of attention vectors which is then used by each decoder in the decoding component. This helps the decoder focus on the appropriate position in the input sequence. The decoder predicts words one word at a time from left to right. Upon prediction of each word, it is again fed into the bottom decoder after converting it into an embedded vector space and adding a positional embedding vector. The decoder's self-attention mechanism works in a slightly different way than the encoder. In the decoder, the self-attention layer is only allowed to look at the words in earlier positions. This is done by masking the words at future positions to *-inf*. Each decoder layer in the decoding component performs the same function. The output of the last decoder layer is then fed into a linear layer and softmax layer. The linear layer outputs a vector having a size equal to the size of the target language vocabulary. Each position in this output vector determines the score of the unique word. This vector of scores is then converted into probabilities by the softmax layer and the position with the highest probability is chosen, and the word associated with it is produced as the output for the particular time step.

4 Experimental Setup

4.1 Dataset

The fundamental requirement for assembling a machine translation system is the availability of parallel corpora of the source and the target language. In this paper, the transformer model is trained on the Hindi-English parallel corpus by the Center for Indian Language Technology (CFILT), IIT Bombay [11]. The training data consists of approximately 1.5 million texts from multiple disciplines while the development and the test set contains data from the news domain. Table 1 provides the details about the number of sentences and the number of unique tokens in English and Hindi that are present in our chosen dataset.

4.2 Data Preprocessing

Data preprocessing is an essential data mining technique that helps clean the data by removing the noise and outliers which can then directly be used by the model for training and testing purposes. Our preprocessing pipeline consists of 3 main steps viz. Data Cleaning, Removal of duplicates and Vocabulary creation. Each step is explained in detail below:

Data Cleaning (Step 1) In this step, the special characters, punctuation and noise characters are removed from both the English and Hindi text corpus. After the elim-

Table 1 Metadata of the dataset

Dataset	# Of sentences	Unique Hindi tokens	Unique English tokens
IITB train	1,267,502	421,050	242,910
IITB dev	483	2479	2405
IITB test	2478	8428	9293

ination of all the noise characters, the empty lines are removed. The resulting text corpus was then converted into lower case and was then fed into the next step to remove the duplicates.

Removal of Duplicates (Step 2) The cleaned and noise-free text corpus obtained as a result of the above step was then used to remove the duplicate records. This resulted in the creation of our training universe containing approximately 1.2 million unique parallel text corpus which was used for creating the vocabulary.

Vocabulary Creation (Step 3) Vocabulary creation is one of the most fundamental step in Neural Machine Translation. The coverage of the vocabulary is a major factor that drives the overall accuracy and the quality of the translation. If the vocabulary under-represents the training data universe, then predicted translation will contain many *unk* tokens thereby reducing the BLEU score drastically. Thus creating a modest-size vocabulary that optimally represents the data is a challenging task. For creating the vocabulary for both Hindi and English language, two approaches are implemented: word level tokenization and subword level tokenization. In the word level tokenization, it first extracted 50,000 most frequently used words from the training set and then added it to the vocabulary. While in the subword level tokenization, Byte Pair Encoding (BPE) is used for creating 50,000 subword tokens which were added in the vocabulary. The evaluation of the performance of our model on both word and subword level tokenization is presented in Sect. 5.

4.3 Back-Translation

Hindi, being a low resource language as compared to its counterpart European languages has made the availability of data quite difficult. Many institutions around the world are creating larger and a more complete text corpus for the low resource languages. To tackle the lack of availability of Hindi-English parallel corpus, back-translation technique is used. Back-translation technique is used for augmenting the training data which leads to increasing the output accuracy of the translation. There is a plethora of monolingual English data available on the internet which can be used to generate text corpus of a low resource language. To generate the additional Hindi-English parallel text corpus, an English to Hindi machine translation system is first trained on our training data and then translated the 3 million WMT14 English monolingual data to generate the corresponding predicted Hindi text corpus.

To observe the effect of back-translation, it first divided the 3 million back-translated parallel records in 4 batches. Then cumulatively added the back-translated records to the original training data in each of these batches. The first batch contains the 0.5 million back-translated records along with the 1.2 million original training data. In the same way, an additional 1 million, 1 million and 0.5 million are added in the second, third and fourth batch respectively. Table 2 summarizes the training data universe for each batch.

Table 2 Training data universe: Batch-wise summary

Batch number	# Of back-translated records added (million)	Total records (million)
Batch 1	0.5	1.7
Batch 2	1.5	2.7
Batch 3	2.5	3.7
Batch 4	3	4.2

4.4 Training Details

After the data preparation and segregation into batches, our transformer model is trained using Opennmt-tf toolkit [12]. For training the model, the NVIDIA Tesla K80 GPU provided by Google Colab [13] is used. For our transformer model, the default 6 layers setting in both encoder and decoder each of which contains 512 hidden units are used. Further the proposed work used the default embedding size of 512 along with 8 headed attention. The batch size is configured to be equal to 64 records and the effective batch size which is defined as the number of training examples consumed in one step to be equal to 384. The model parameters are optimized using the LazyAdam optimizer. The model was trained on 10 different configurations 5 each for word and subword level tokenization till convergence or till 70,000 steps at max (hard stop). The GPU run-time provided on Google Colab resulted in a training duration of approximately 20–24 h for each configuration.

5 Results

The quality of translation of our model are trained on the test set using the Bilingual Evaluation Understudy (BLEU) score [14] and the Rank-based Intuitive Bilingual Evaluation (RIBES) score [15]. For depicting the performance of subword level tokenization, where it divides the test set into 2 subsets. The first set (Set-1) consists of sentences whose words are present in the vocabulary generated from word level tokenization. This set consists of 1694 sentences. The second set (Set-2) is the complete test set consisting of 2478 sentences.

In Table 3, after adding the first batch of 0.5 M parallel back-translated records with the original training data, the BLEU score increased by 3.79 and with the subsequent addition of other 2 batches the BLEU and the RIBES score reached a maximum of 24.79 and 0.741 respectively. However, when the 4th batch of 0.5 M back-translated records was added with the previous batches, the scores decreased by a small margin indicating convergence with respect to the addition of back-translated data.

Similar to the results obtained with word level tokenization, in Table 4, after adding the first batch of back-translated records the BLEU score increases by 4.78 and with

Table 3 Results of word level tokenization (Set-1)

Model ID	Model	BLEU	RIBES
1	Transformer	18.76	0.699708
2	Transformer with Batch 1	22.55	0.730440
3	Transformer with Batch 2	23.95	0.735804
4	Transformer with Batch 3	24.79	0.741369
5	Transformer with Batch 4	24.68	0.740567

Table 4 Results of subword level tokenization (Set-1)

Model ID	Model	BLEU	RIBES
6	Transformer	19.10	0.695566
7	Transformer with Batch 1	23.98	0.733614
8	Transformer with Batch 2	25.44	0.737078
9	Transformer with Batch 3	25.87	0.739192
10	Transformer with Batch 4	25.74	0.742397

Table 5 Results for the transformer with Batch 4 model on Set-2

Model ID	Tokenization	BLEU	RIBES
5	Word level	21.22	0.728683
10	Subword level	24.53	0.735781

the subsequent addition of other 2 batches the BLEU score reached a maximum of 25.87. After adding the 4th batch, the BLEU scored decreased by 0.13 however the RIBES score increased by 0.003.

When compared with word level tokenization, subword level tokenization achieves a better BLEU score which can be attributed to the fact that it has the advantage of not having an out-of-vocabulary case and also to learn better embeddings for rare words since rare words can enhance the learning from its subwords that occur in other words. This fact is further strengthened in Table 5 which shows the BLEU and RIBES score for the Transformer with Batch4 model using word and subword level tokenization on Set-2. The decrease in the BLEU and RIBES score as compared to Tables 3 and 4 is due to the fact that the Set-2 consists of additional sentences as compared to Set-1 which contain rare words that are not included in the vocabulary for word level tokenization. When subword level tokenization is used, the model performs reasonably well even in the presence of rare words which is not the case for word level tokenization.

6 Conclusion

The transformer model has displayed promising results for neural machine translation involving low resource languages as well. It is observed that after adding the back-translated records the performance was certainly improved, however when the amount of generated data increases beyond a certain level, there is no further improvement in the performance. Using a combination of the transformer model, back-translation technique and a subword tokenization method like BPE, a BLEU score of 24.53 which is the state-of-the-art on this dataset to the best of our knowledge has been achieved. The future research directions on the proposed work will try to incorporate state-of-the-art Natural Language Processing models like BERT [16] into NMT to further improve the quality of translation.

References

1. Saini S, Sahula V (2018) Neural machine translation for English to Hindi. In: Fourth international conference on information retrieval and knowledge management (CAMP). IEEE, pp 1–6
2. Rule-based machine translation. https://en.wikipedia.org/wiki/Rule-based_machine_translation
3. Yang S, Wang Y, Chu X (2020) A survey of deep learning techniques for neural machine translation
4. Luong M, Sutskever I, Le Q, Vinyals O, Zaremba W (2014) Addressing the rare word problem in neural machine translation
5. Tacorda A, Ignacio M, Oco N, Roxas R (2017) Controlling byte pair encoding for neural machine translation. In: International conference on Asian language processing (IALP) 2017. IEEE, pp 168–171
6. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A, Kaiser Ł (2017) Attention is all you need. In: Advances in neural information processing systems, pp 5998–6008
7. Sennrich R, Haddow B, Birch A (2015) Improving neural machine translation models with monolingual data
8. Sennrich R, Haddow B, Birch A (2015) Neural machine translation of rare words with subword units
9. Goyal V, Sharma D (2019) LTRC-MT simple & effective Hindi-English neural machine translation systems at WAT 2019. In: Proceedings of the 6th workshop on Asian translation 2019, pp. 137–140
10. The illustrated transformer. http://jalammar.github.io/illustrated-transformer
11. Kunchukuttan A, Mehta P, Bhattacharyya P (2017) The IIT Bombay English-Hindi parallel corpus
12. Klein G, Kim Y, Deng Y, Senellart J, Rush A (2017) OpenNMT: open-source toolkit for neural machine translation. In: Proceedings of ACL, system demonstrations 2017. Association for Computational Linguistics, pp 67–72
13. Google Colab. http://colab.research.google.com/
14. Papineni K, Roukos S, Ward T, Zhu W (2002) BLEU: a method for automatic evaluation of machine translation. In: Proceedings of the 40th annual meeting. Association for Computational Linguistics 2002, pp 311–318

15. Isozaki H, Hirao T, Duh K, Sudoh K, Tsukada H (2010) Automatic evaluation of translation quality for distant language pairs. In: Conference on empirical methods in natural language processing, vol 2010, pp 944–952
16. Devlin J, Chang M, Lee K, Toutanova K (2018) Bert: pre-training of deep bidirectional transformers for language understanding

CNN-Based Alphabet Identification and Sorting Robotic Arm

Saleh Al-Faraj, Mustafa Al-Bahrani, Saeed Al-Ghamdi, Marwan Rafie, Abul Bashar, and Ghazanfar Latif

Abstract Automated identification of objects and sorting them based on specified criteria is a crucial problem which is encountered by various manufacturing companies. To this end, they employ automated robots which need to perform these tasks with high accuracy. In this paper, a CNN-based machine learning system is proposed, designed, and implemented to accurately identify objects marked with English alphabets and sort them in a correct order based on the input given by the user. It consists of a hardware module which incorporates a robotic arm controlled by a Raspberry Pi microcontroller. The software module is based on a CNN-based image identifier model trained on an indigenous dataset consisting of 3898 images of English alphabets rotated at random angles. The experimental results demonstrate training and validation accuracies of 99.06% and 98.79%, respectively, based on the model trained over 120 epochs. Furthermore, our system was able to successfully sort and arrange the alphabets with the desired accuracy.

Keywords Alphabet sorting · Robotic arm · Image recognition · Convolutional neural networks · Robotic automation

---

S. Al-Faraj · M. Al-Bahrani · A. Bashar (✉) · G. Latif
Computer Engineering Department, Prince Mohammad Bin Fahd University, Khobar, Saudi Arabia
e-mail: abashar@pmu.edu.sa

S. Al-Faraj
e-mail: 201501220@pmu.edu.sa

M. Al-Bahrani
e-mail: 201601053@pmu.edu.sa

G. Latif
e-mail: glatif@pmu.edu.sa

S. Al-Ghamdi · M. Rafie
Computer Science Department, Prince Mohammad Bin Fahd University, Khobar, Saudi Arabia
e-mail: 201601001@pmu.edu.sa

M. Rafie
e-mail: 201600035@pmu.edu.sa

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_26

1 Introduction

Generally speaking, sorting is such a common problem observed in various applications and researchers have tried to find fast and efficient ways to sort. Warehouse systems utilize the need for robots to manage their inventory and sort goods using automated systems. There have been many solutions toward automated sorting problem which will be presented in the next section. Alphabetical sorting is one problem that has not been given due attention and is a challenge the industry faces today. The main idea of this paper is to propose an automated robotic arm for identifying and sorting objects which have English alphabets written on them. Essentially, in our approach, a robotic arm is set up, which has the means to read images through a camera and identify the alphabet with the aid of computer vision algorithms. To make the system more interactive, the user inputs a word and based on this the robotic arm is supposed to pick the alphabets in the correct order and build the word with the correct spelling (i.e., sort the alphabets correctly). Robotic arm in general is one of the best options suitable for arranging and sorting problems. The proposed system utilizes a Raspberry Pi microcontroller to control the robotic arm movement, while the alphabet identification is performed with the help of a CNN-based machine learning model which is trained on our generated dataset.

The rest of this article is organized as follows. Section 2 discusses the background of this domain, while Sect. 3 presents the architectural details of the proposed system. Section 4 provides the results achieved and the related discussion. Section 5 concludes the paper with the main contributions and future recommendations.

2 Background

To design our system, the review of the literature is done for the automated sorting problem and found papers where the approach is either arranging objects or recognizing handwritten alphabets using a model which is pretrained dataset (such as the MNIST dataset) [1–4]. Digital image processing becomes important field with the emergence of new machine learning and deep learning techniques with the availability of high computational resources for the recognition, classification, and segmentation in different areas such as medical diagnosis from images [5], sign language recognition for disabled people [6, 7], traffic signs recognition [8], image enhancement [9, 10] and assisted living for visually impaired persons [11, 12]. Unable to find any research work that would arrange alphabets at a rotated angle which is scattered in a region of interest, it is believed that this is a new take on arranging scattered alphabets that are rotated at some random angle.

Table 1 provides a summary of the related work in this domain, where each approach is compared to our approach in terms of the key idea, their benefits, and drawbacks. This was also a discussion that wanted to know how beneficial our research work can be used as a real-world application. It is understood that some

Table 1 Comparison of different existing techniques with our proposed solution

References	Main idea	Benefits	Drawbacks
[13]	Recognizes alphabets based on their color. The arm detects the position of the alphabets from the image. The arm sorts the alphabets according to their color	Sorting colored objects from the panel and put them in the right boxes	There are many different items with the same color
[14]	Robotic arm sorts alphabets based on speech recognition. Alphabets are recognized based on RGB color and shape from the camera images	It avoids mistake and can carry heavyweight	The sound may face a problem since there are different accents diversity
[15]	A webcam is used to captures colored object cubes. Based on the color of the cubes, the robotics arm will place them into different cups	It works well for the manufacturer that uses color in organizing products	Based on the color of the image so leads to miss recognition if colors change
[16]	The camera image is processed using GNU Octave to determine the color and the shape	It is low cost and can sort colored objects	The objects around it can lead to wrong shape detection measurements
[17]	A handwritten optical character recognition (OCR) from the camera images to control the robotic arm	It can be used to scan handwritten notes and similar texts	Possibility for the wrong prediction if it is written by different persons
Proposed alphabet sorting robotic arm	The camera will take a picture of the workspace, and the alphabets are recognized using CNN that will send the position to the Robotics arm and arrange them accordingly	Using CNN as classifier gives high accuracy	It will only work for English Alphabets

industries have an arranging issue, where some packages would be labeled, and they would need to arrange them according to some criteria. One example application would be a library, as people come and borrow books. Once they return it, all these books would be added in some cart, shuffled, with labels that indicate which shelf should be placed alphabetically and it is a hassle for a librarian to sift through these books and find the "correct" book with the specified labeled to be added to the shelf. This is one possible application, however, there could be many other applications to use our proposed solution.

3 Proposed System Design

The proposed system has two major modules, namely the hardware and the software. The hardware module consists of the robotic arm, Raspberry Pi, and the servo controller. The software module is composed of the OCR system and the CNN-based image identifier and alphabet classifier. Their details are now presented below.

3.1 New Alphabets Dataset

A new data is prepared which has 3898 images A to Z capital alphabets as shown in Table 2. Firstly, the data processing packages were imported and then the images were captured from the images dataset. The sample of captured images is shown in Fig. 1.

Here, it can be seen that the alphabets are randomly rotated a certain angle. Next, they imported the machine learning packages, loaded the processed data, and split them into testing data and training data. The test data had 20% of the total number of images while the training data had 80% of the images. Later, a predefined model with four layers is created. The next step was to train the CNN model.

The model was then validated and found to be fit for testing. After that, came up with the trained and validate model which have been used to recognize the randomly arranged alphabets. The above-mentioned steps of alphabets image recognition are depicted in Fig. 2.

Table 2 Newly build alphabets dataset

A	190	F	128	K	155	P	127	U	131	Z	157
B	157	G	152	L	155	Q	153	V	149	Total	
C	158	H	155	M	155	R	137	W	152	3898	
D	126	I	156	N	157	S	140	X	149		
E	156	J	149	O	155	T	145	Y	154		

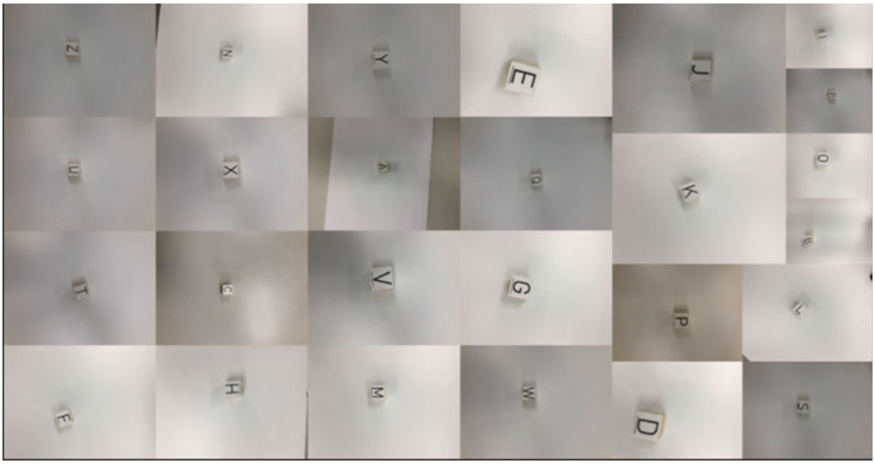


Fig. 1 Sample images of the newly built dataset

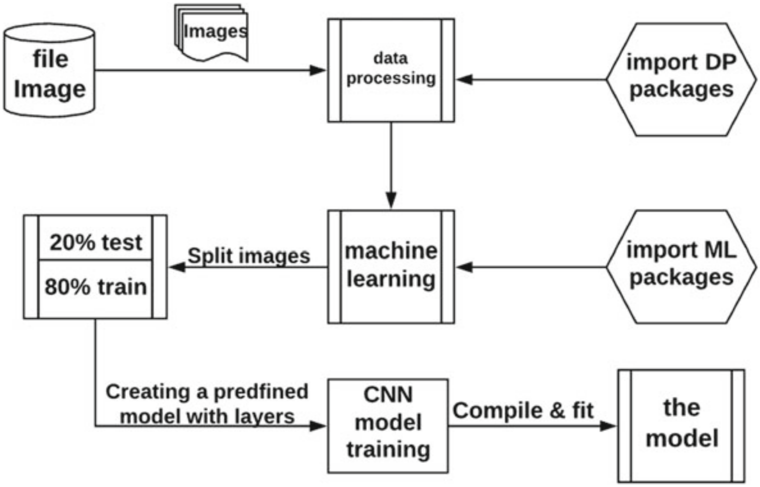


Fig. 2 Process of alphabets image recognition

3.2 Alphabets Image Processing

Now, the process of the alphabet OCR is described which is shown in Fig. 3. The camera takes the picture of the workspace which is saved as an image in a local file. Then, the alphabets in the image are recognized using optical character recognition (OCR). OCR is a technology that translates an image to be recognized by a machine. It distinguishes printed or handwritten text within digital images. Also, it is mostly used to scan documents. The software module first imports the packages and reads

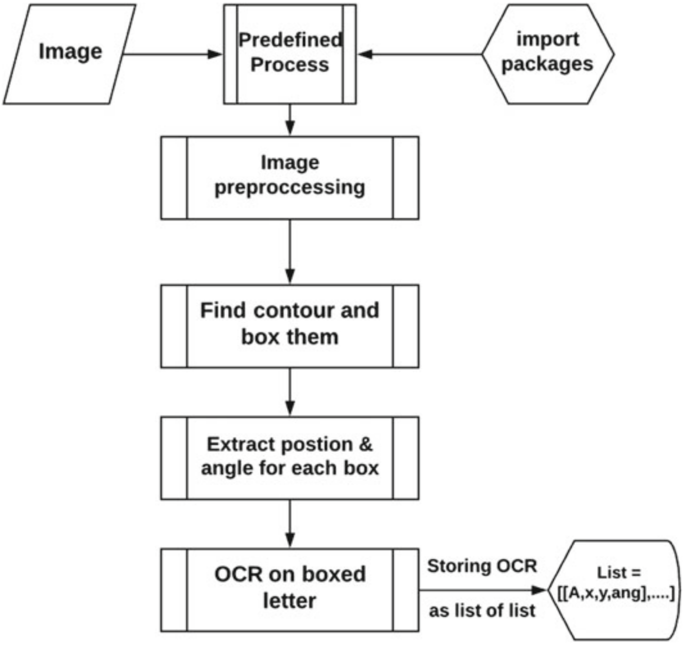


Fig. 3 Workflow of the proposed alphabets OCR

the input image. Then, it converts the image to grayscale and thresholds the image using the OpenCV library. After the recognition of the alphabet through CNN, the image is processed to find the contour and then they are boxed. After that, it extracts the position and the angle for each box. The midpoint for each box is then found, to have the arm fixed at the center of each box. Then in a repeated manner, it will perform the OCR on boxed alphabets to get the output as a list. The list will contain [*alphabets*, *x* and *y* coordinates, *rotation*]. The *alphabet* is one of the 26 possible English alphabets, the *x* and *y* coordinates represent the position on the workspace and the *rotation.* is the angle at which the alphabet is arranged. There are three main functions of the alphabets image processing, which are described below. They are preprocessing step, contour identification step, and the storage step.

Preprocessing: There is a need to preprocess the input which got as an image of the alphabets available on the workspace. The image is converted into black and white. Then threshold it using binary inversion. Also, there is a need to dilate the alphabets so that the OCR can better detect the desired alphabets.

Contour: After preprocessing our image, each alphabet has to be boded to figure out the rotated angle of each alphabet, locate the *x*-position and *y*-position of the alphabets, and perform an OCR to detect the alphabets shown in the box.

Data Storage: Once the OCR completes detecting each alphabet, it will store them as a list where each list contains four data variables: the alphabets detected, *x*-position, *y*-position and the angle. This array will then be passed to the robotic arm to loop through the alphabets and arrange them accordingly.

3.3 Alphabets Recognition Through CNN

Alphabets recognition was done with the convolutional neural network-based machine learning approach. Figure 4 shows the proposed CNN model used for image recognition where the image size is 48 × 48 and it passes through convolution 2D layer consisting of 48 filters with a kernel size of 3 × 3 and ReLU activation function [18, 19]. Then it passes through a max pooling 2D layer with 32 filters having a kernel size of 2 × 2 which returns the important features present in the image. This will result in the reduction of the image size. Then, it goes through another convolution layer of 24 filters with kernel size 3 × 3 and having a ReLU activation function. Later the image passes through a third convolution layer of 12 filters having a kernel size 3 × 3 and consisting of ReLU activation function. Then, having the classification layer that flattens the matrix and converts the vectors into a fully connected layer. Finally, the softmax function is used to classify the image into alphabets based on the predicted probability distribution value.

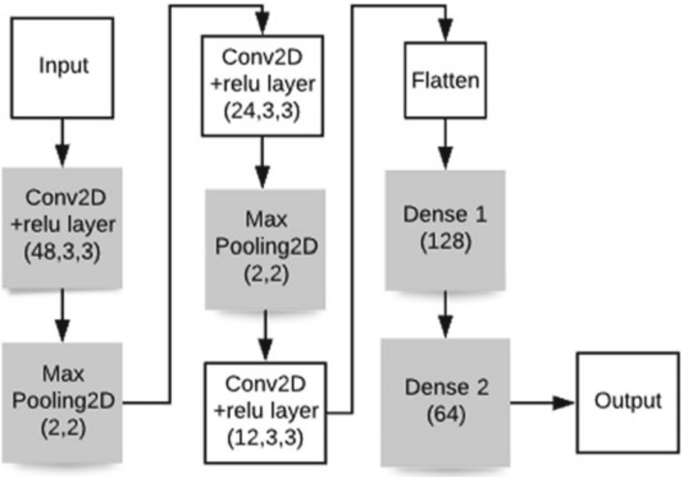
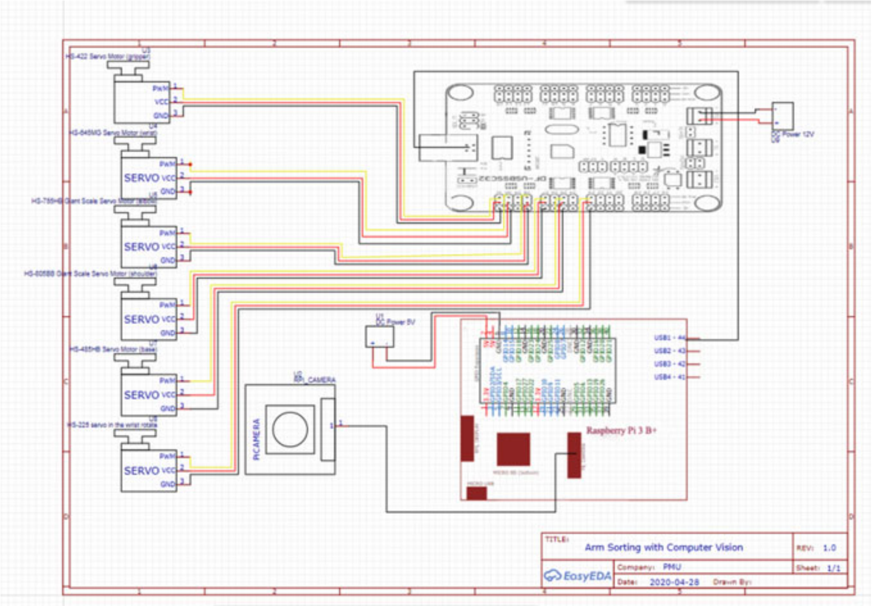

Input
Conv2D +relu layer (48,3,3)
Max Pooling2D (2,2)
Conv2D +relu layer (24,3,3)
Max Pooling2D (2,2)
Conv2D +relu layer (12,3,3)
Flatten
Dense 1 (128)
Dense 2 (64)
Output

Fig. 4 Proposed CNN architecture

3.4 Hardware Design and Implementation

In the proposed system, different hardware components are integrated with the robotic arm which includes Raspberry Pi, digital camera, and RC servo motors. Figure 5 shows the circuit diagram of the hardware components. The hardware components are Raspberry Pi model B+, SSC-32U USB servo controller, six different types of HS servo motors, Pi camera, and two DC power supplies. The six HS servo motors are connected into six different channels in SSC-32U USB servo controller. Each channel contains three inputs which are pulse width modulation (PWM) pin, VCC pin, and ground pin. These inputs are arranged from top to bottom for each channel. Each servo motor has three outputs which are pulse width modulation (PWM) pin, VCC pin, and ground pin. SSC-32U USB servo controller requires a 12 V power supply connected to the VS2 because the channels that are used are connected to VS2. However, Raspberry Pi3 model B+ requires a 5 V power supply. The positive side of the DC power supply is connected to pin 4 in the Raspberry Pi, and the negative side of the DC power supply is connected to pin 6 in the Raspberry Pi. Raspberry Pi is connected with the SSC-32U USB servo controller through the USB cable. There is a special port in the Raspberry Pi (number 24) which is used to connect the Pi camera. The type of cable that is used to connect Raspberry Pi and Pi camera is CSI.

Robotic Arm: AL5D-PLTW arm is part of Lynxmotion's collection of AL5 robotic arm as shown in Fig. 6. This robotic arm has four degrees of freedom (4-DOF).



U3
HS-422 Servo Motor (gripper)
U4
HS-645MG Servo Motor (wrist)
U5
HS-755HB Giant Scale Servo Motor (elbow)
U6
HS-805BB Giant Scale Servo Motor (shoulder)
U7
HS-485HB Servo Motor (base)
U8
HS-225 servo in the wrist rotate
SERVO
PWM
VCC
GND
1
2
3
U1
DC Power 5V
U2
RPI_CAMERA
PICAMERA
DC Power 12V
USB1 - 44
USB2 - 43
USB3 - 42
USB4 - 41
Raspberry Pi 3 B+
TITLE:
Arm Sorting with Computer Vision
REV: 1.0
EasyEDA
Company: PMU
Sheet: 1/1
Date: 2020-04-28
Drawn By:

Fig. 5 Circuit diagram of the hardware components

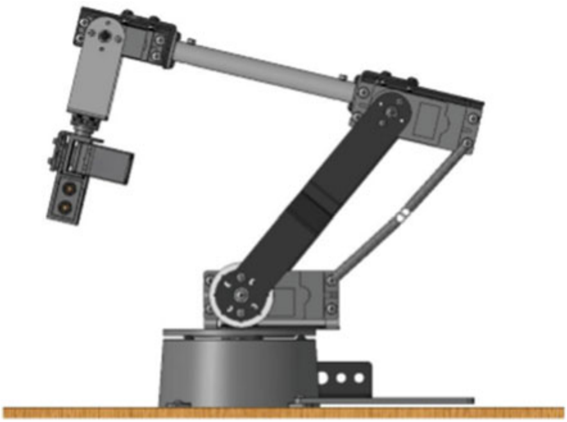


Fig. 6 AL5D robotic arm

The AL5D robotic arm can perform repeatable movements with high accuracy [20]. One important feature is that the arm can move at high speeds with precise positional placements [21]. Dimensions of the AL5D robotic arm: shoulder to elbow: 14.605 cm. Elbow to wrist: 18.7325 cm. Wrist to tip of gripper: 8.5725 cm. Height: 18.415 cm. Height (reaching up): 48.26 cm. Median forward reach: approx. 26.035 cm. Gripper opening: 3.175 cm. Weight: 0.878 kg. Range of motion per axis: 180°.

Servo Controller: SSC-32U is a dedicated servo controller, the core of R/C servo controller is an ATmega328p chip which has a Harvard architecture with an 8-bit RISC processor core as shown in Fig. 7 [22]. The servo controller was not supposed to be programmed but was meant to receive and execute commands sent to it from an external system such as a computer or microcontroller like Raspberry Pi. The R/C servo controller has many features including control up to 32 servo motors, USB, and serial input.

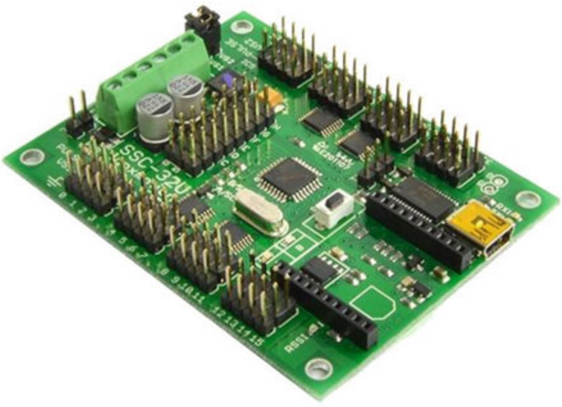


Fig. 7 SSC-32U servo controller

Raspberry Pi Microcontroller: It is a small credit card-sized computer. It is not only limited to perform routine home automation tasks, but can also be used for various other applications such as home entertainment, a video game console, or anything that is programmable [23]. Latest models of Raspberry Pi have in the better processing capabilities, new features, such as a wireless and Bluetooth chip. For our research work, it is found useful based on its relation to some of the system requirements desired, such as moving the robot arm to arrange the alphabets and executing a computer vision libraries which can recognize the alphabets. Therefore, the Raspberry Pi 3 B+ model is decided to use and installed the Raspbian OS to initiate our research work.

Workspace Setup: The workspace is made up of a square wooden board. It consists of two crossed planks, which are supported on top by three wooden planks as shown in Fig. 8. These crossed planks support the camera which is placed in the middle to capture the wooden board (workspace). The size of our wooden board is roughly around $20 \times 20\,\text{in}^2$. The workspace contains the alphabets which are scattered around having random angular rotations. The robotic arm is placed at the other end of the wooden board, so it will only be able to reach the alphabets of the workspace in front of it. The distance between the camera and the wooden board is around 64 in. The alphabets are written on cubic-shaped wooden boxes. Each side of the cubic wooden box is exactly 1 in., so the robotic arm grabber (end effector) will be able to grasp it and place it at the desired location on the workspace. The position information will be given as an input to the robotic arm from the output of the OCR module explained earlier.

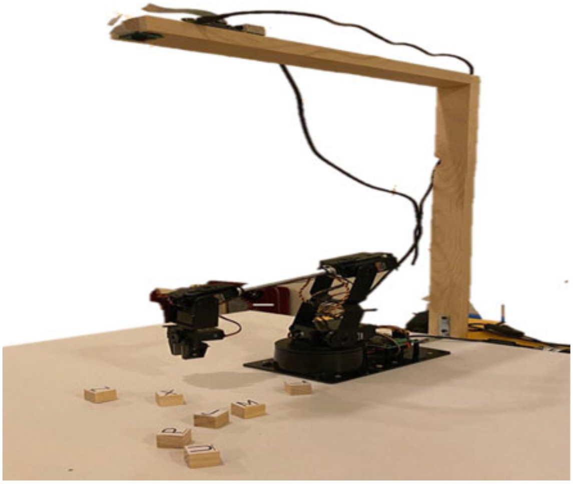


Fig. 8 Workspace setup for the robotic arm system

4 Experimental Results

In this section, the results are now presented from the experiments which were performed on the proposed system. Even though our system has two modules (hardware and software) focused on the results related to the software module, the hardware module was able to work properly based on the control signals provided to the Raspberry Pi controller from the CNN-based image classification module.

Two important accuracy measures which focused on were training accuracy and validation accuracy and also measured training loss and validation loss. These are important measures when evaluating the performance of CNN-based image classifier. The number of epochs was the independent variable which was used in this study. The number of epochs was varied from 1 to 120 and its effect on the four performance metrics is presented in Table 3. As it is expected, the accuracies increase with the number of epochs, whereas the loss functions decrease. Figure 9 shows that the training accuracy steadily increases from 21% to a maximum value of 99%, whereas

Table 3 Experimental results for alphabets recognition using CNN with different epochs

Epochs	Train Acc (%)	Train loss	Val. Acc (%)	Val. loss
1	21.46	2.7689	16.82	2.9277
15	65.83	1.0079	27.19	3.1293
30	77.81	0.6876	31.49	3.3836
60	87.50	0.4657	78.06	0.8680
120	99.06	0.2444	98.79	0.2408

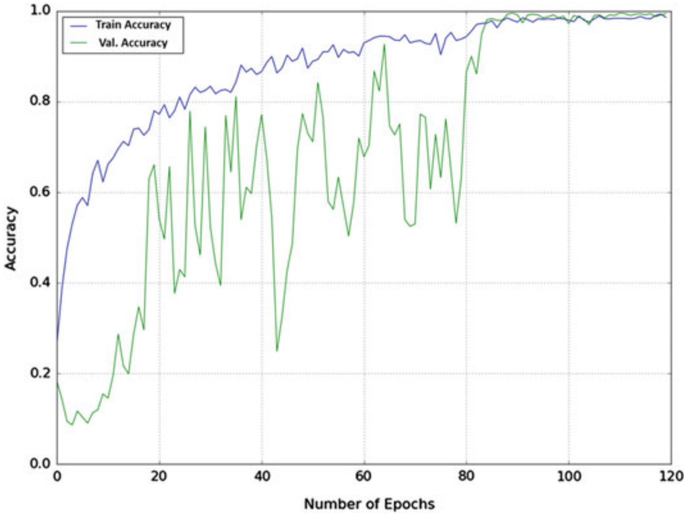


Fig. 9 Accuracy for alphabets recognition using CNN

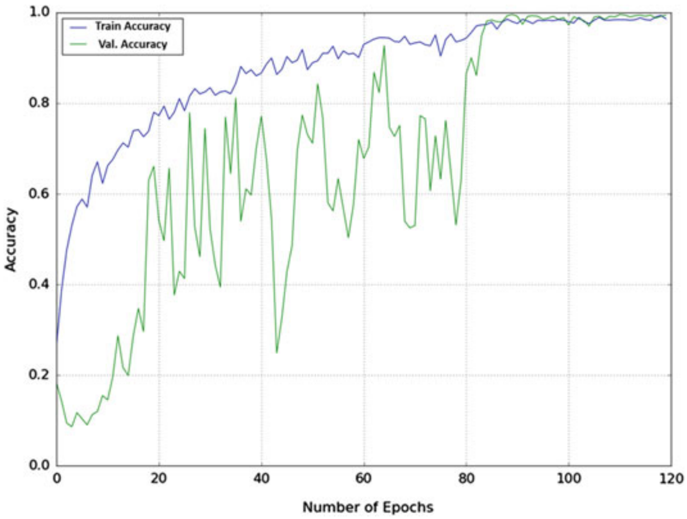


Fig. 10 Loss measurements for alphabets recognition

the validation accuracy initially has a fluctuated behavior and later on stabilizes to a maximum value of about 98.8%. Figure 10 provides a loss measurements for the recognition of the alphabet. As can be observed, both the training loss and the validation loss steadily decrease from a value of 2.77 to 2.93, respectively. At an epoch value of 120, the loss values reach a minimum value of 0.24 for both cases. By training the CNN model with 120 epochs, it found out that our system was able to correctly identify and classify the alphabets, as the accuracy was close to 99%. Further, it is observed that the robotic arm was able to correctly arrange the alphabets in the desired order according to the word which was given as an input to the system. Based on a variety of words, our system was able to demonstrate the achievement of the proposed objective which was English alphabet sorting.

5 Conclusions

In this paper, a deep learning-based CNN model has been proposed, designed, and implemented for recognizing and sorting English alphabets. Hardware and software modules are combined to provide a solution for arranging alphabets based on the desired word given as an input by the user. The hardware part of the system consists of a Raspberry Pi 3 Model B+, SSC-32U servo controller, and RC servo. The software part of the system was programmed in Python and using a convolutional neural network (CNN) machine learning model. The CNN model was trained on a dataset consisting of 3898 images of the 26 English alphabets randomly rotated at certain angles. The experimental results show that our model had a training and validation accuracies of 99.06% and 98.79%, respectively. As a part of future work, a system is planned to implement for Arabic alphabets. The changes that need to be implemented is the recognition of 28 Arabic alphabets and for this, the required layers to be used in the CNN model have to be studied. Another extension would be to speed up the training process by optimizing the various CNN layers and to reduce the identification time through faster processing microcontrollers.

Acknowledgements Our sincere thanks to the College of Computer Engineering and Science at Prince Mohammad Bin Fahd University for providing the resources and facilities for conducting this research.

References

1. Latif G, Alghazo J, Alzubaidi L, Naseer MM, Alghazo Y (2018) Deep convolutional neural network for recognition of unified multi-language handwritten numerals. In: 2018 IEEE 2nd ınternational workshop on Arabic and derived script analysis and recognition (ASAR). IEEE, pp 90–95
2. Al-Hmouz A, Latif G, Alghazo J, Al-Hmouz R (2020) Enhanced numeral recognition for handwritten multi-language numerals using fuzzy set-based decision mechanism. Int J Mach Learn Comput 10(1)
3. Alghazo JM, Latif G, Alzubaidi L, Elhassan A (2019) Multi-language handwritten digits recognition based on novel structural features. J Imaging Sci Technol 63(2):20502–20511
4. Alghazo JM, Latif G, Elhassan A, Alzubaidi L, Al-Hmouz A, Al-Hmouz R (2017) An online numeral recognition system using ımproved structural features–a unified method for handwritten Arabic and Persian numerals. J Telecommun Electron Comput Eng (JTEC) 9(2–10):33–40
5. Latif G, Iskandar DA, Alghazo JM, Mohammad N (2018) Enhanced MR image classification using hybrid statistical and wavelets features. IEEE Access 7:9634–9644
6. Latif G, Mohammad N, AlKhalaf R, AlKhalaf R, Alghazo J, Khan M (2020) An automatic Arabic sign language recognition system based on deep CNN: an assistive system for the deaf and hard of hearing. Int J Comput Dig Syst 9(4):715–724
7. Latif G, Mohammad N, Alghazo J, AlKhalaf R, AlKhalaf R (2019) ArASL: Arabic alphabets sign language dataset. Data Brief 23:103777
8. Alghmgham DA, Latif G, Alghazo J, Alzubaidi L (2019) Autonomous traffic sign (ATSR) detection and recognition using deep CNN. Procedia Comput Sci 163:266–274
9. Khan AH, Al-Asad JF, Latif G (2017) Speckle suppression in medical ultrasound images through Schur decomposition. IET Image Proc 12(3):307–313
10. Latif G, Iskandar DA, Alghazo J, Butt M, Khan AH (2018) Deep CNN based MR image denoising for tumor segmentation using watershed transform. Int J Eng Technol 7(2.3):37–42
11. AlSaid H, AlKhatib L, AlOraidh A, AlHaidar S, Bashar A (2019)Deep learning assisted smart glasses as educational aid for visually challenged students. In: Proceedings of ınternational conference on new trends in computing sciences (ICTCS 2019), Amman, Jordan, pp 1–6
12. AlZamil M, AlBugmi R, AlOtaibi S, AlAnazi G, AlZubaidi L, Bashar A (2020) COMPASS: IPS-based navigation system for visually ımpaired students. In: Proceedings of 9th IEEE ınternational conference on communication systems and network technologies (CSNT 2020), Gwalior, India, pp 161–166
13. Jia Y, Yang G, Saniie J (2017)Real-time color-based sorting robotic arm system. In: 2017 IEEE ınternational conference on electro ınformation technology (EIT), Lincoln, NE, pp 354–358. https://doi.org/10.1109/EIT.2017.8053385

14. AlSalman Z, AlSomali N, AlSayari S, Bashar A (2018) Speech driven robotic arm for sorting objects based on colors and shapes. In: 2018 3rd ınternational conference on ınventive computation technologies (ICICT), Coimbatore, India, pp 6–11. https://doi.org/10.1109/ICICT43934.2018.9034306
15. Szabó R, Lie I (2012)Automated colored object sorting application for robotic arms. In: 2012 10th ınternational symposium on electronics and telecommunications, Timisoara, pp 95–98. https://doi.org/10.1109/ISETC.2012.640811
16. Pereira V, Fernandes VA, Sequeira J (2014)Low cost object sorting robotic arm using Raspberry Pi. In: 2014 IEEE global humanitarian technology conference—South Asia Satellite (GHTC-SAS), Trivandrum, pp 1–6. https://doi.org/10.1109/GHTC-SAS.2014.6967550
17. Sarma P, Chourasia CK, Barman M (2019)Handwritten Assamese character recognition. In: 2019 IEEE 5th ınternational conference for convergence in technology (I2CT), Bombay, India, pp 1–6. https://doi.org/10.1109/I2CT45611.2019.9033603
18. Butt MM, Latif G, Iskandar DA, Alghazo J, Khan AH (2019) Multi-channel convolutions neural network based diabetic retinopathy detection from fundus images. Procedia Comput Sci 163:283–291
19. Shaikh E, Mohiuddin I, Manzoor A, Latif G, Mohammad N (2019) Automated grading for handwritten answer sheets using convolutional neural networks. In: 2019 2nd ınternational conference on new trends in computing sciences (ICTCS). IEEE, pp 1–6
20. LynxmotionLynxmotion SSC-32U USB Servo Controller Board. Retrieved from: https://www.robotshop.com/media/files/pdf2/lynxmotion_ssc-32u_usb_user_guide.pdf32u_usb_user_guide.pdf. Last accessed on 28 Apr 2019
21. Montiel-Vázquez EC, Torres-Rosique AE, Park G, Garcıa-Cavazos I, González-Flores U, y Mejia-Rosete RS, González-Hernández HG MOBMA: gesture driven mobile manipulator
22. Saleh T, Khan MR (2019) Hexapod robot for autonomous machining. In: IOP conference series: materials science and engineering, vol 488, no 1. IOP Publishing, p 012003
23. Agrawal N, Singhal S (2015) Smart drip irrigation system using raspberry pi and arduino. In: International conference on computing, communication and automation. IEEE, pp 928–932

Lung Cancer Detection from LDCT Images Using Deep Convolutional Neural Networks

Shahad Alghamdi, Mariam Alabkari, Fatima Aljishi, Ghazanfar Latif, and Abul Bashar

Abstract Lung cancer is second cancer common to men and women as well as it is one of the world's highest cause of death. Reports in recent years have shown that standard X-rays are not effective in diagnosing lung cancer. It has clinically established that low-dose computed tomography (LDCT)-based diagnosis helps to decreases mortality from lung cancer by 20% relative to normal chest X-rays. Deep learning is considered as one of the most beneficial techniques for lung cancer diagnosis. This technique used in many fields, including healthcare, which helps to facilitate complex tasks, analyze medical images, promote reliable diagnosis, and improve diagnostic accuracy. One of the deep learning algorithms is the convolutional neural network (CNN) and in this paper, different deep CNN based models are proposed for lungs cancer detection. The experiments are performed using dataset acquired from Data Science Bowl 2017 (KDSB17). The dataset consists of 6691 LDCT lung images. For testing the efficiency of the model, the accuracy is reckoned, which represents 91.75%. However, due to the sensitivity of this process, other techniques are also used to assess the model's performance including specificity, sensitivity, recall, precision, and f1-score.

Keywords Deep learning (DL) · Convolutional neural networks (CNNs) · Low-dose computed tomography (LDCT) · Lung image database consortium (LIDC)

1 Introduction

Lung cancer or lung carcinoma, regardless of race, is the principal cause of death from cancer among both men and women. More people die each year from lung cancer than from other popular cancers [1]. It has the lowest several rates among colorectal cancer (65%), prostate cancer (99%), and breast cancer (89%). In 2018, 1.76 million deaths have been estimated [2, 3]. For several decades, lung cancer was

---

S. Alghamdi · M. Alabkari · F. Aljishi · G. Latif (✉) · A. Bashar
College of Computer Engineering and Sciences, Prince Mohammad Bin Fahd University, Al Khobar, Saudi Arabia
e-mail: glatif@pmu.edu.sa

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_27

the most prevalent and deadliest form of cancer. It is the most common cancer among men and is the third most common cancer among women worldwide as two million new cases have been reported in 2018 [4]. Statistics show that over 50% of patients with lung cancer die within one year of diagnosis. In 2014, lung carcinoma was Saudi men's fourth most frequent cancer and Saudi women's 7th frequent cancer [5].

Medical imaging generates an immense volume of data, and thousands of images are involved in each medical study. Deep learning is used in healthcare to improve diagnosis accuracy, resolution, and promote reliable and fast diagnosis [6–9]. It can extract and process medical images at a speed and scale that exceed human capabilities and analyze more efficiently. Even though lung cancer is the deadliest type of cancer, it is highly curable if diagnosed early. Catching lung cancer before spreading can add years to human life. It can increase the possibility of survival for five years or more by 55% [1], so it is important to invest in a system that assists in early lung cancer detection. England statistics show that 88% of lung cancer patients who diagnosed at the first stage survived for at least one year compared to 19% of those diagnosed at the fourth stage [2]. Even though it is the number one cause of death from cancer, less money is spent on research into lung cancer as compared with other rising cancers.

Deep learning algorithms consists of several layers which extract higher-level characteristics from the raw input. This consists of several layers for extracting characteristics of higher levels from the raw inputs [10, 11]. Each layer in the network transforms its input into a more complex and abstract representation. The first representational layers learn to detect simple feature filters such as edges and corners while the middle representational layers learn more complex feature detection filters such as part of an object. The last layer will learn to recognize the full object. The research puts in a profoundly convolutional neural network model (CNN) to diagnose lung cancer patients into two classes: have cancer and does not have cancer. The network comprises two convolution layers, two max-pooling layers, two drops out layers, two fully connected layers, and a flattened layer.

In recent years, doctors found that normal X-rays are not appropriate for the diagnosis of lung cancer. They found that low-dose computed tomography (LDCT) decreases mortality from lung cancer by 20% relative to normal chest X-rays. comparison to regular chest radiography [2]. LDCT can detect pulmonary cancer early on, which not measurable with a standard X-ray. Thus, a data set consisting of 6691 LDCT images have been used for the experiments in this research [12].

2 Literature Review

Serj, Lavi, Hoff, and Valls gave a new deep convolutional neural network. (dCNN) model for learning high-level image representation to achieve highly accurate results with low variance using low-dose computed tomography (LDCT) images [13]. The researchers' goal was building a binary classification model that learns discriminant compact features at the beginning of the network to detect lung cancer. The researchers proposed a new deep convolutional neural network architecture consists of three convolution layers, two max-pooling layers, a full-connected layer, and a binary soft-max layer. The proposed model commences with several sequential convolution layers to generate high-order convolutional features. The researchers used a data set from the Data Science Bowl 2017 (KDSB17) to confirm the model results. The dataset consists of 63,890 (LDCT) images of cancer patients and 171,345 images of non-cancers. The researchers divided the data set into testing, training, and cross-validate set; 50% for the training and 25% for the validation and the rest for the testing. The researchers used cross-entropy as a loss function to maximize the probability of having cancer by maximizing the multinomial logistic regression objective. The results of the model were impressive. The performance of the model is measured using specificity (0.991), sensitivity (0.87), and F1 score (0.95).

Chon and Balachandar have used the deep convolutional neural network for lung cancer detection of CT scans [14]. They used a data set from Date Science Bowl 2017 along with updated U-Net that trained on data set LUNA16, which represents CT scans with marked nodules. At first, they started to pass the CT scans into the 3D CNNs directly for classification, which gave them a poor result. Then, to insert only the regions of interest into the 3D CNNs, they had to perform further pre-processing. For this to be the case, U-Net that trained on the LUNA16 data set was used (CT scans with labeled nodules) for identification in CT scans of nodule candidates. This process produced many false positive predictions, so they used the CTs scans to specify where the nodules located as determined by the U-Net outputs were fed into 3D convolutional neural networks. That ultimately helped to identify CT scans for lung cancer either as positive or negative.

Several kinds of research have applied lung cancer diagnosis to manipulating images and machine learning approaches. Makaju, Prasad, Singh, and Alsadoon had two models, their best model was not reliable and failed to identify the findings of cancer found in the nodules [15]. Therefore, they also introduced a new method for detecting cancer nodule from the CT scan image using the watershed segmentation to identify along with SVM to classify a nodule as malignant or benign. This model has identified 92% accuracy of cancer, this is better than the previous model which had 86.6% accuracy. Regarding their data set, they used actual CT scans from the Lung Image Database Consortium (LIDC) archive of patients. This collection of pulmonary cancer CT images for computer-aided diagnostic methods for the identification and treatment of lung cancer, which was introduced by the National Cancer Institute. The dataset was composed of 1018 cases which supported 7 research centers and 8 medical imaging firms. Images were in DICOM format with 512 * 512 pixels in size, and as DICOM format was hard to process; those images have been converted to JPE grayscale images the aid of MicroDicom software, which converts the DICOM CT scan images to JPEG format.

Sharma and Jindal obtained their CT images from NIH/NCI Lung Image Database Consortium (LIDC) data set that is provided for research purposes [16]. They obtained an automated CAD program to diagnose lung cancer early they achieved that by examining the CT images in multiple steps, their method started by extracting pulmonary areas from the CT picture using multiple image processing strategy, including bit image slicing, erosion, and Weiner filter. Instead of using the thresholding technique, they used the bit plane slicing technique that was used to convert the CT image into a binary image during the first stage of the extraction process. This strategy is faster and user autonomous compared to the thresholding technique. Afterwards, the lung regions extracted were segmented using region-widening segmentation algorithms. Lastly, the field has been used to classify cancerous regions and to obtain an objective result with a high sensitivity level of 90%, with a fair number of false positives per picture with 0.05 false positives per picture.

Medial image processing gained more importance in last decade due to the availability of high computational power and advancement in the imaging techniques such as medical image enhancement [17], noise removal [18], medical image classification [19], and medical image segmentation [20].

3 Methodology

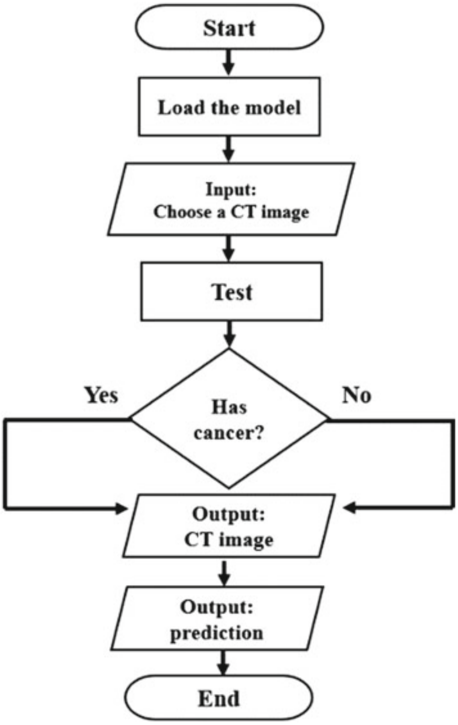
In this section, the models that have been tried will be discussed. The deep convolutional neural networks (dCNN) is generated by an input layer, hidden layers, and an output layer. It composes of two major parts: feature extraction and classification. The first layers learn basic detection filters while the complexity increases in the middle layers. The biggest advantage of dCNN, the developer is not expected to extract features manually from the image. During the training, the network learns to extract features. The classification decision will be guided in the last layers based on the features extracted from the preceding layers. Figure 1 represents the process of lung cancer detection system.

3.1 Classification Through Deep CNN

Medical images are often corrupted by noise, lighting, and affected by artifacts; this could have an impact on model accuracy. So, to limit these phenomena, the data set was divided by 255. 75% of the images were used for the training, and 25% for the testing. To balance the classes in each set, the data has been divided based on the labels, and the random state is set to 42. Before deciding our model, 4 deep convolutional neural network architectures were tried then compared the results and chosen the most accurate, which is the fourth model. Figure 2 shows the workflow of the deep convolutional neural network for lungs cancer detection.

The first model consists of three sequential convolution layers, two max-pooling layers, two fully connected layers, and a flattened layer. Firstly, the network begins with an input layer which is the first convolution layer; the first layer will take the image with an input size of 120 × 120 pixels. It consists of 50 filters and convolution kernel of 11 × 11. Secondly, another convolution layer has been added. The second layer consists of 120 filters and convolution kernel of 5 × 5. Thirdly, a max-pooling

Fig. 1 Process of lung cancer detection system



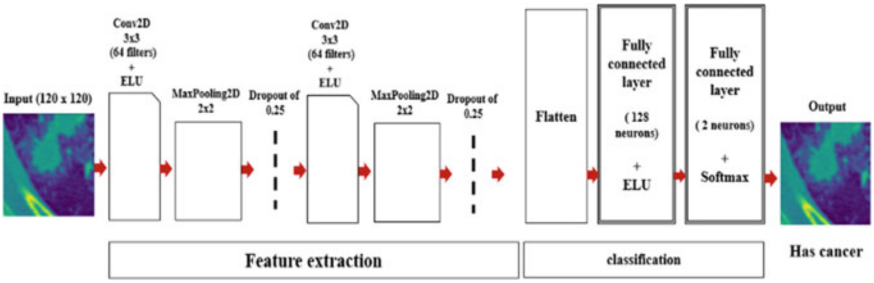


Fig. 2 Workflow of the deep convolutional neural network for lungs cancer detection

layer has been added with a pool size of 2 × 2 and strides of 2. Fourthly, a third convolution layer with 120 filters and convolution kernel of 3 × 3 is added. Fifthly, a max-pooling layer has been added with a pool size of 2 × 2 and strides of 2. RELU is used as an activation function in all the convolution layers. This model did not get the best results, this might be because the flattened layer is added after the first dense layer.

In the second model, the first convolution layer is the input layer. The input layer consists of 32 filters and with the convolution kernel of 2 × 2. After that, the batch normalization technique is used to speed up the training, reduce overfitting, and to put all our data on the same scale. Similar to batch normalization, the drop-out layer will help in reducing overfitting. After that, a convolution layer consists of 64 filters is added with the convolution kernel of 2 × 2. After that, the batch normalization is used and the drop-out layer again. After the feature extraction and normalization, the output of the final convolution layer will be the input of the flatten layer. The flatten layer will convert the data into a one-dimensional array and pass it to the first fully connected layer. After that, a drop-out layer is added between the dense layers because they have the largest number of parameters and could cause overfitting. Finally, the output layer will pass its output to the last fully connected layer, and the classification decision will be made.

The third model consists of an input layer that takes an image input (120, 120, 1); this convolution layer applies 50 kernels (filters) of size 50 with kernel size (11, 11). Then, the process of batch normalization is used to improve training and prevent overfitting. As with batch normalization, the drop-out layer helps to reduce overfitting. After that, a max-pooling layer is used with pool size 2 × 2 and strides of 2. A layer of convolution that consists of 80 filters with 3 × 3 kernel size is applied. Following that, batch normalization is used and the drop-out layer again. Afterwards, max-pooling layer with pool size 2 × 2 and strides of 2 is added. Finally, the output of the last layer will pass through the flatten layer which will convert the data into a 1-dimensional array and pass it to the first fully connected layer.

Because having one input and one output, the network begins with a sequential convolution layer consisting of 64 filters with a convolution kernel of 3 × 3. The layer uses the exponential linear unit (ELU) as an activation function. In addition, it takes the LDCT image of an input size of 120 × 120 pixels. Secondly, a max-pooling layer with a 2 × 2 pool size has been added. This layer was used to minimize the number of parameters along with minimizing the spatial size of the representation. Therefore, the computational cost and overfitting will be reduced [18]. Similarly, a dropout of 0.25 has been added to prevent the model from overfitting. Fourthly, a convolution layer consists of 64 filters a convolution kernel of 3 × 3 is added. The fourth layer uses ELU as an activation function. Fifthly, a max-pooling layer with a 2 × 2 pool size is added. Sixthly, a dropout of 0.25 is added. After that, two flatten layers is added to convert the pooled feature map to a single column to input it to the fully connected layer. Finally, two fully connected layers are added to take the results of the convolution and pooling process and use them to drive a classification decision. As shown in Fig. 1, the first fully connected layer uses ELU as an activation function while the second fully connected layer uses SoftMax. The SoftMax is usually applied in the last layer of the neural networks instead of using RELU, Tanh, or Sigmoid. It is used because it converts the input into values between 0 or 1, so they can be interpreted as probabilities [19]. After constructing the layers, the model is trained for 100 epochs. In addition, 10% of the training set is used for validation. Because Adam optimizer is used, the accuracy increased in each epoch.

3.2 Experimental Data

The data is acquired from the cancer imaging archive (TCIA) as a file of un-labeled scans. The file contains two datasets, ‘ct_slices’ which had 6691 slices of the original scans to help with the sharpness of the image when resizing it [12]. All the images are sorted in a hierarchical data format (HDF5). The second dataset ‘slice_class’ had the labels of each slice. The labels used were 0 and 1 to indicate the presence of cancer. The 6691 images consist of 2526 cancerous images, and 4165 non-cancerous images. Sample images of the CT lung cancer are shown in Fig. 3.

The original size of the images was 64 × 64 pixels, yet all the images were resized to 120 × 120 pixels by using resize function from cv2 library. This helped to obtain clear slices and detect cancer cells accurately while using the model. NumPy library which contains reshape function is used to reshape the images from [6691, 64, 64] to the new shape [6691, 120, 120, 1]. Reshaped it from a three-dimensional array to a four-dimensional array to use it with the model. 6691 represents the number of the images, 120 represents the height and width of the images, and 1 represents the number of channels in the images as all the images are in grayscale. Training the model by using grayscale images may increase the model performance because the model will focus on the shape of the images rather than the colors of each image [21]. Also, used to categorical function to convert the labels to a matrix to use it with the model since having two classes 1: cancerous and 0: non-cancerous. The dataset

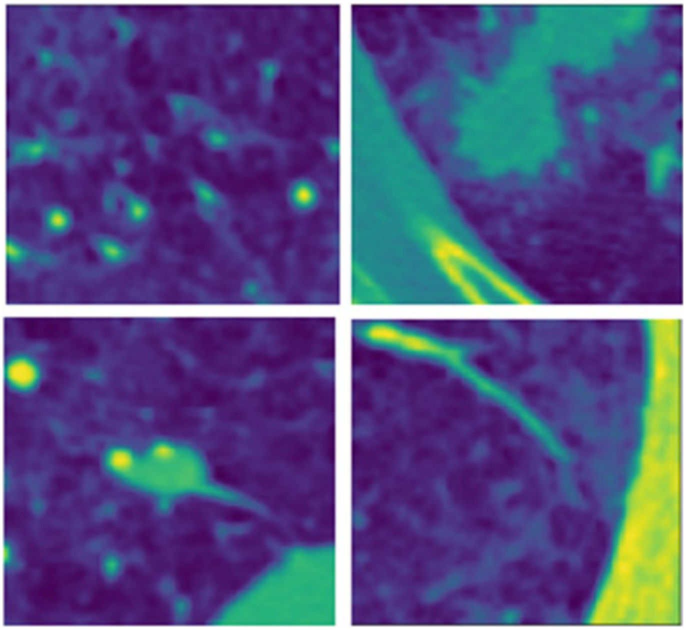


Fig. 3 Sample of low-dose computed tomography (LDCT) lung images

has been divided randomly into three categories: 65% of the images used for the training, 10% used to validate the model during the training process, and 25% used to test the model.

4 Experimental Results

The dataset has been divided randomly into three categories: 65% of the images used for the training, 10% used to validate the model during the training process, and 25% used to test the model. The dataset has been divided based on the labels to guarantee that each category includes images of both classes 0: non-cancer and 1: cancer. To achieve the objective of this research, four different experiments are conducted along with different features for the best results. In the first experiment, the model is trained with 200 epochs, which means 200 training cycles. In addition, the model was complicated, because it contained many layers with a high number of filters and big kernel size. Thus, it was the cause of overfitting. In the second and third experiments, tried to improve the model by decreasing the number of epochs from 200 to 100 cycles. The number of layers used is decreased along with reducing the number of kernels and kernel size. Different parameters are also applied to avoid the overfitting and increase the model performance. This helped to avoid overfitting; however, it affected the accuracy of the model. An accuracy of 91.75% is achieved in the fourth experiment, as long with high model performance.

The accuracy of the four experiments is calculated to evaluate the overall efficacy of the classification process. The accuracy is simply representing the percentage of the predictions that the model has been getting correct. Formally, the accuracy determined by using the formula below for binary classification:

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

Due to the sensitivity of this process, several techniques is used to evaluate our model's efficiency, including specificity, sensitivity, recall, precision, and f1-score. This will help us to calculate the percentage of the right and wrong predictions, which will help to avoid having wrong predictions since the software could have an impact on society as it could be used in medical diagnosis. A wrong diagnosis will affect individuals' health. As shown in the table below, a recall is determined showing the right positive rating average from all the real positive ratings. In addition, the precision is calculated from the cases predicted as positive that represents the rate of correct positive classification. The f1-score helped us to combine the recall and precision values to balance and see the overall results.

In medical image analysis, model performance usually measured using sensitivity and specificity, which is the percentage of true positive and true negative. The geometric mean rate shows a combination of sensitivity and specificity in a single matric.

Table 1 Comparison of the accuracies of different proposed DCNN Models

Epochs	Accuracy (%)
Proposed dCNN model M1	83.02
Proposed dCNN model M2	77.23
Proposed dCNN model M3	71.73
Proposed dCNN model M4	91.75

Different models were implemented to compare their performances and choose the best accuracy among them where the fourth model had the highest accuracy of 91.75% as shown in Table 1. The variance of the accuracy was because of the changes made in each model including, the type of layers added, number of kernels, kernel size, number of epochs (training cycles), and type of activation functions. In addition, more parameters have been added to boost model efficiencies such as kernel initializer, Adam optimizer, Elu activation, and some padding.

A sequential model is created to add multiple layers to the model, where each layer has only one input and one output. The proposed deep convolutional neural network model consists of two convolution layers, two max-pooling layers, two drop-out layers, a flatten layer, and two fully connected layers. Multiple layers are applied and repeated some of the layers multiple times for higher performance, along with extract more features from the images which may help in the classification process.

The two convolution layers applied 64 kernels of the size of 3 × 3. The 3 × 3 filter is the smallest and the most commonly used. In the fourth model, the size of the filter is tried to reduce from 11 × 11 to 3 × 3. This is because a smaller filter size of an odd number is preferred over large filter size of an even number. This will help simplify the image processing for better performance of the convolution layers, along with the number of kernels (filters) used to extract useful features from the images.

Elu activation function is used which more likely to converge cost to zero faster. In addition, Elu tends to produce more accurate results, and it combines the good features of ReLU and Leaky ReLU. Moreover, Elu does not have the main problems of ReLU. The same padding in both convolution layers is applied, which helped to keep the dimensions of the output the same as its input. That means the convolution layers output size is the same as the input. For instance, the input shape of the first convolution layer was (120, 120, 1) and the output shape was also (120, 120, 64). This helped to not reduce the features of the images and lose any important feature, which may help in the classification process.

Two drop-out layers are added after applied the Elu activation function inside the convolution layers; this technique helped us to avoid having overfitting. Each hidden unit (neuron) is set to 0 with a probability of 0.25 at passing 0.25. This means that there is a change of 25% forcing the neuron output to 0. The detailed proposed model's performance analysis is presented in Table 2 which compares different proposed models experimental results based on PrecisionRecall F1-score Specificity.

Adam optimizer is used from Keras library which helped to improve the model and increase the accuracy during the training process. This will guarantee that a

Table 2 In-depth proposed model's performance analysis

#		Precision	Recall	F1-score	Specificity
M1	Non-cancerous	0.87	0.86	0.86	0.79
	Cancerous	0.77	0.79	0.78	0.86
M2	Non-cancerous	0.74	0.98	0.84	0.42
	Cancerous	0.94	0.42	0.58	0.98
M3	Non-cancerous	0.70	0.97	0.81	0.30
	Cancerous	0.86	0.30	0.44	0.97
M4	Non-cancerous	0.92	0.95	0.93	0.86
	Cancerous	0.92	0.86	0.89	0.95

higher accuracy will get at each training cycle. In addition, Keras initializer is used to pass initializers to the layers. In addition, the number of dense is increased, which represents the number of neurons the full connection layer will connect to from 10 to 128 connections; this helped to get a perfect test error.

Lastly, the epoch number was reduced from 200 to 100 to prevent overfitting, because the first model is trained with 200 epochs, and the model started to memorize the images which gave us 1 accuracy during all the training cycles. During this process, the model has trained with only 6691 (LDCT) images, so no need to have a high number of epochs unless the number of images used is increased.

5 Conclusion

Studies have shown that computed tomography (LDCT) based diagnosis helps to decreases mortality from lung cancer by 20% relative to normal chest X-rays. The deep learning technique helped to analyze the (LDCT) medical images and promoted reliable diagnosis and improved the diagnosis accuracy. The paper proposed profoundly convolutional neural network architecture for binary lung cancer classification. Convolutional neural network (CNN) is one of the profound learning algorithms, which used to achieve high classification accuracy for tasks involving medical images. The convolutional neural network (CNN) model consists of two convolution layers, two max-pooling layers, two drop-out layers, a flatten layer, and two fully connected layers. CNN layers were used to extract features from the images, which permit the model to be trained using these features to be able to make the prediction. The models presented, predicted, and classified the cancerous and non-cancerous slices with 91.75% accuracy. However, due to the sensitivity of this process, a lot of methods are used to assess our model's performance, including specificity, sensitivity, recall, precision, and f1-score. These techniques helped to calculate the percentage of true positive, true negative, false positive, and false negative levels, which will help to avoid having wrong predictions since the software could have an impact on society as it could be used in medical diagnosis. This process will help to promote a reliable diagnosis along with improving the diagnosis accuracy.

References

1. Desai A, Gyawali B (2020) Fall in US cancer death rates: time to pop the champagne? EClinicalMedicine 19
2. Rindi G, Klimstra DS, Abedi-Ardekani B, Asa SL, Bosman FT, Brambilla E et al (2018) A common classification framework for neuroendocrine neoplasms: an International Agency for Research on Cancer (IARC) and World Health Organization (WHO) expert consensus proposal. Mod Pathol 31(12):1770–1786
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68(6):394–424
4. Latif G, Butt MM, Khan AH, Butt O, Iskandar DA (2017) Multiclass brain Glioma tumor classification using block-based 3D Wavelet features of MR images. In: 2017 4th ınternational conference on electrical and electronic engineering (ICEEE). IEEE, pp 333–337
5. Jazieh AR, AlGhamdi M, AlGhanem S, AlGarni M, AlKattan K, AlRujaib M et al (2018) Saudi lung cancer prevention and screening guidelines. Annals Thor Med 13(4):198
6. Kim M, Yun J, Cho Y, Shin K, Jang R, Bae HJ, Kim N (2019) Deep learning in medical imaging. Neurospine 16(4):657
7. Butt MM, Latif G, Iskandar DA, Alghazo J, Khan AH (2019) Multi-channel convolutions neural network based diabetic retinopathy detection from fundus images. Procedia Comput Sci 163:283–291
8. Latif G, Iskandar DA, Alghazo J, Butt MM (2020) Brain MR ımage classification for glioma tumor detection using deep convolutional neural network features Current Med Imag
9. Latif G, Iskandar DA, Alghazo J, Butt M, Khan AH (2018) Deep CNN based MR image denoising for tumor segmentation using watershed transform. Int J Eng Technol 7(2.3):37–42
10. Alghmgham DA, Latif G, Alghazo J, Alzubaidi L (2019) Autonomous traffic sign (ATSR) detection and recognition using deep CNN. Procedia Comput Sci 163:266–274
11. Latif G, Alghazo J, Alzubaidi L, Naseer MM, Alghazo Y (2018) Deep convolutional neural network for recognition of unified multi-language handwritten numerals. In: 2018 IEEE 2nd ınternational workshop on Arabic and derived script analysis and recognition (ASAR). IEEE, pp 90–95
12. Jadhav S (2020) Lung cancer detection using classification algorithms (Doctoral dissertation, Dublin, National College of Ireland)
13. Frank M, Drikakis D, Charissis V (2020) Machine-learning methods for computational science and engineering. Computation 8(1):15
14. Chon A, Balachandar N, Lu P (2017) Deep convolutional neural networks for lung cancer detection. Standford University
15. Makaju S, Prasad PWC, Alsadoon A, Singh AK, Elchouemi A (2018) Lung cancer detection using CT scan images. Procedia Comput Sci 125:107–114
16. Sharma D, Jindal G (2011) Identifying lung cancer using image processing techniques. In: International conference on computational techniques and artificial ıntelligence (ICCTAI), vol 17, pp 872–880
17. Khan AH, Al-Asad JF, Latif G (2017) Speckle suppression in medical ultrasound images through Schur decomposition. IET Image Proc 12(3):307–313
18. Al-Asad JF, Khan AH, Latif G, Hajji W (2019) QR based despeckling approach for medical ultrasound images. Current Med Imag 15(7):679–688
19. Latif G, Iskandar DA, Alghazo JM, Mohammad N (2018) Enhanced MR image classification using hybrid statistical and wavelets features. IEEE Access 7:9634–9644

20. Latif G, Iskandar DA, Jaffar A, Butt MM (2017) Multimodal brain tumor segmentation using neighboring image features. J Telecommun Electron Comput Eng (JTEC) 9(2–9):37–42
21. Joulin A, Cissé M, Grangier D, Jégou H (2017) Efficient softmax approximation for gpus. In: International conference on machine learning, pp 1302–1310

Global Integration and Distribution of Data Through Machine Learning for COVID-19

E. Laxmi Lydia, Jose Moses Gummadi, Chinmaya Ranjan Pattanaik,
G. Jaya Suma, A. Krishna Mohan, and Ravuri Daniel

Abstract COVID-19 is referred to as a broad disaster struck out in the society as a challenge. The large quantity of statistical information requires machine tools to improve the knowledge and accelerate COVID-19 forecast, analysis, and its corresponding remedial measures. But, to evade global hazards in these applications, open research will be made mandatory. This article uses machine learning model to integrate COVID-19 data and distribute the data globally. Machine Learning (ML) solutions that rely on COVID-19 data use the SIR model and logistic regression model to analyze how the pandemic cycle propagates all around the inhabitants. The figures of the SIR model concentrate on time-driven case scenarios to predict the behavior of infection, whereas the cumulative cases have become more reliant on significant-data plots.

Keywords COVID-19 · Machine Learning (ML) · SIR model · Logistic Regression (LR)

E. Laxmi Lydia (✉)
Department of Computer Science and Engineering, Vignan's Institute of Information Technology (A), Visakhapatnam, Andhra Pradesh, India
e-mail: elaxmi2002@yahoo.com

J. M. Gummadi
Department of CSE, VFSTR (Deemed To Be University), Guntur, India
e-mail: josemoses@gmail.com

C. R. Pattanaik
Department Computer Science and Engineering, Ajay Binay Institute of Technology, Cuttack, Odisha, India
e-mail: chinmaya.pattnaik@rediffmail.com

G. Jaya Suma
JNTUK-UCEV, Vizianagaram, India

A. Krishna Mohan
Department of Computer Science and Engineering, JNTUK, Vizianagaram, Andhra Pradesh, India

R. Daniel
Department of Computer Science and Engineering, Bapatla Engineering College (Autonomous), Bapatla, India
e-mail: danielravuri@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_28

1 Introduction

The first coronavirus case has been originated from the city of Wuhan, China and infected many humans across the globe since December 31, 2019. Organizations from China and World Health Organization (WHO) declared the existence of novel coronavirus and also reported the related diseases, especially SARS-CoV-2 and COVID-19. 213 countries got infected to date by crossing 2.4 million virus cases worldwide [1] with death cases 4, 18,135 till June 10, 2020. COVID-19 was diagnosed and pointed as a great danger to global health. The unusual drastic increase of virus and its roots escalation was underlined unknowingly within the society. Global distribution of experimental COVID-19 data overture assuring tools to flap against the disease. More than 12,400 editorials were printed in the preceding five months [2], and the information regarding COVID-19 thousands of patients were recorded and broadcasted [3]. The maximum investigation of the data was gathered with blood sampled and interpreted. Certainly, the standard mechanism to grab success over this pandemic circumstance experiences two common asserts. The demand for proper input and substantial expertise is necessary. Secondly, it should maintain timely assumptions. These two assets were meet by machine learning pathologies [4, 5] comprising of disease infections [6]. The following are the two statistical approaches using machine learning for the COVID-19 pandemic and also the process to share data and uphold global collaboration.

2 Literature Survey

Machine learning phase medication for COVID-19 is an interpretive framework that usually takes a few assertions into account at a time. For comparative purposes, out of over 1,200 pre-clinical studies documented for symptoms for COVID-19, the majority of the population continues to focus is on a single medicine or a few drugs, hand-selected based on different relations [7]. The machine learning process can be broadened to choose a set of alternative antiviral; several patterns of DNAs and/or protein molecule that can anticipate interrelationships among medicines and the virus, including potential SARS-CoV-2 drug sites, to allow impressive treatment methods for candidates [8, 9]. ML has indeed been employed similarly in certain contagious disorders [10]; for instance, a deep neural network has been specially managed to monitor the operation of the *Escherichia coli* [11] activity of more than 100 million organic compounds. In a certain case, a broad spectrum of vaccination patients could also be tested individually in order, for obvious reasons by conveying the dramatic drop protein S observing a SARS-CoV-2 infection [12], to develop successful immune reactions. Such prospective pathways do not, however, mask the complexities of pharmacological ML-based work. Initially, ML still could not stimulate fundamental biology, and it a remarkable problem even without the prediction of protein [13].

Therefore, there is indeed mandatory time duration in the particular instance of vaccines. Secondly, a massive essential problem is the desire to avoid appropriate laboratory tests: The latest hydroxychloroquine-based therapeutic research [14] has indeed documented functioning with very small samples, which don't use sufficient model or obfuscating the eligibility criteria. With ML algorithms, this risk could significantly increase. Methodologies, along with deep neural networks, are "broad sense classifiers." They could be organized to match each primary goal on a dataset by memorizing medications of all patients. Analytical techniques are sometimes assessed only in a compelling way by evaluating their ability to estimate an independent test set.

3 Methodology

3.1 Alleviate the Workload of Medical Experts Through Machine Learning

Although conventional statistical techniques might yield the very first required responses through a disaster, they also enable significant information systems that are distinctly insufficient during this scenario. Health care institutions easily became overloaded, and the capacity for statistical analyses was restricted mostly beyond the amount of effort expected, related to clinical investigation. Predictive models may reduce the period deemed necessary through analytical research and assist clinicians via artificial intelligence practitioners. Diagnostic imaging analyses, for instance, demonstrate that chest verified tomography (CT) diagnostic tests will be utilized to pinpoint COVID-19 infections [15, 16]. Conversely, this research method involves a skilled and experienced radiologist to evaluate each sample, who otherwise would be operating across the front lines. ML can help solve this task: for the diagnosis of 14 distinct lung infections, subsequently supervised classifiers selected around a set of data of 400,000 chest X-Rays managed to achieve a normalized area under its patient operating characteristic curve (true positive rate vs. false positive) of 94% [17].

Also, observational research focused on several hundred chest cartridge diagnostic tests indicates whether COVID-19 can be diagnosed often with ML software automatically [18]. Consequently, clinical images ML is currently restricted to extremely minor comparisons to predict or diagnose COVID-19; many such investigations, therefore, have little influence over the many confusions (e.g., age, corpulence) here which methodologies can find from images in the chest. One suitable strategy is to pre-train machine learning models from bigger image datasets and develops learning computational attributes that might be used to promote COVID-19 image training. In recent years, this technique has been used consistently in computer vision to produce successful outcomes with very few case studies [19].

3.2 Essential Data Integrity

ML is pertinent to speed up the analytics of challenging and complicated statistical techniques, including large genomic or medical imagery data samples, although standard statistical analyses are assimilated to many clinical and epidemic issues. Cumulative ML is also planned instead of replacing traditional diagnosis, forecast, and treatment methods. Nevertheless, the environmental effect of ML is currently constrained by two major issues. Initially, it is notorious that ML algorithms are not easy to understand. While visualization tools may outline the set of variables that led technology to find a certain estimation.

Systemic perceptions (e.g., measuring tool, personage, etc.) can eventually influence ML. Distinctive analytic activities are therefore made to create normal concern in scientific articles as well as in clinics concerning ML results. Furthermore, the absence of huge corporate repositories for public health, clinic, imagery, and genetics leads each organization to localize their analytical platform within its limited data, which noticeably restricts results interpretation. Although this problem does not relate to ML, sophisticated methodologies should lead to disparate datasets (1) accessing at every new treatment of de-anonymized statistical data and (2) the creation of large samples. The ISARIC project intends to choose a robust and popular clinical database of COVID-19 patients [20]. The International Severe Acute Infection Consortium (ISARIC) data-sharing protocols are signed by several other managements to make ensure data which is extensively and easily shared [21, 22] such that innovative concepts are updated, and this is only accomplished in part, making it very hard to use the frequent data captured throughout the deadly virus.

This will be important to facilitate successful initiatives across cultures and various types of health care facilities [6], and perhaps, even directly relies on the scale, functionality, and interpretability of these databases. The interactive exchange of clinical databases calls for careful management of regulatory problems and privacy. Throughout the event that many government agencies do not ordinarily operate, rapid resolution of such problems can be exceedingly difficult. Yet machine learning cannot maintain its commitment which will counteract the virus once to overcome many challenges.

3.3 SIR Model with Machine Learning Regressions

The main objective was to develop a predictive framework for the interpretation of the vital factor influencing the data transfer of COVID-19. SIR is a current idea that takes into account the population from one of the countries listed:

(a) Susceptible (S). The patient has not developed symptoms, but that can be contaminated by affected individuals
(b) Infected (I) with bacteria. This individual does have the infection.

(c) Recovered (R) or Dead. Anyone of two individual lives seems to be the epidemic: whether another individual persists or the patient has become resistant to this disease or has passed away.

Multiple implementations of this model are available, concerning birth and death (SIRD including demographic).

Due to the advancement in COVID-19 aids to be concerned everywhere to enhance the future predictions, furthermore, it is important to discover that individuals gain immunity (a prolonged probability of losing immunity and of returning COVID-19 up in a particular variability like influenza virus) even though there is no transformation from being regained to any of the two States. The differential equations governing the device are as follows:

$$\mathrm{d}S/\mathrm{d}t = -\frac{\beta \mathrm{SI}}{N}$$

$$\mathrm{d}I/\mathrm{d}t = \beta \mathrm{SIN} - \gamma I$$

$$\mathrm{d}R/\mathrm{d}t = \gamma I$$

Here, β denotes the contagion rate of the pathogen, and γ denotes the recovery rate of the epidemic.

The strategy of the SIR has been applied across several aspects: inside a medium-range exponential function of the differential equations and also in the virtual community (graph) further with dynamics. For consistency purposes, the first alternative approach was preferred and implemented for the differential equation system using a numerical method (Runge–Kutta). Consequently, it is necessary to classify the key parameters to get the progression of the disease and notify the rk4 approach.

3.4 Procedure for Standard Logistic Regression Model

(a) Initially choose appropriate features.
(b) Filter data from a certain period.
(c) Perform log transformations to reported cases and deaths.
(d) Substitute logarithm infinity by 0. With the asymptotic logarithm behavior of log(0), a reverse (exponential) transformation has been achieved.
(e) Split the data for training, validating, and testing.
(f) Prediction results.
(g) Finally, report the outcome within that appropriate sequence and execute an exponential transformation to the reverse log.

4 Results

The rate of deaths in India is predicted by the cumulative number of confirmed COVID-19 contexts. Results were carried out using Kaggle data collection, i.e., population by country 2020.csv. Figure 1 shows the SIR model; Fig. 2 shows logistic regression for Indian infected cases.

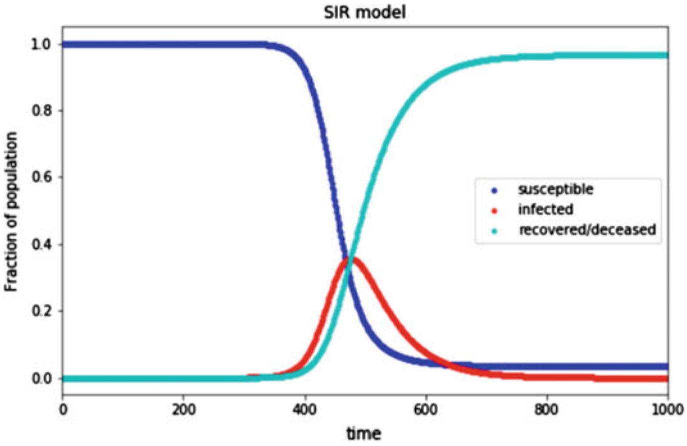


Fig. 1 Plot for SIR model over a fraction of the population

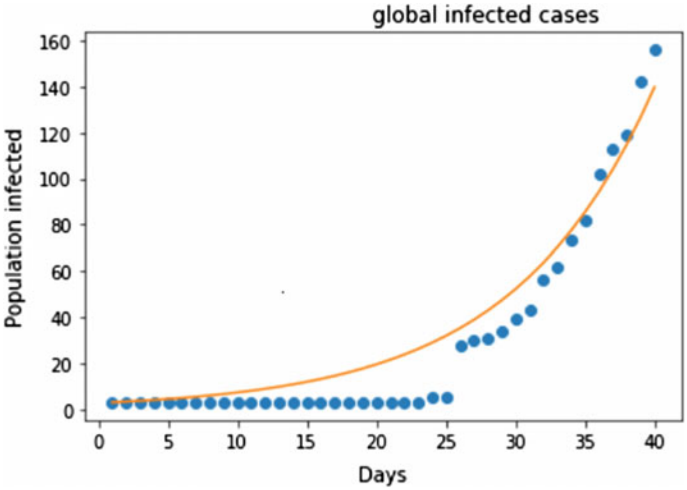


Fig. 2 Logistic regression plot for global infected cases

5 Conclusion

The COVID-19 contagious disease spreading the virus all over but rather it is highly improbable to be the last, all over the world. Worldwide health statements from World Health Organization association and teamwork with several other evidence-based platforms also happen to a single platform to cooperate as a unified system. However, this result's high quality as it entirely depends on machine learning. Its efficiency and consistency are often more focused on developing worldwide partnerships. Transmitting significant information which always accelerates innovation and promoting optimistic interventions. It has been observed that with the spread of the virus, the susceptible percentage of the community will gradually fall to zero, and also from a significant period, where the number of infected cases tends to increase. However, it starts decreasing as people started to recover from the disease.

References

1. Dong E, Du H, Gardner L (2020) Lancet Infect Dis 20:533–534
2. Dimensions COVID-19 publications, data sets, clinical trails. Figshare https://dimensions.figshare.com/articles/Dimensions_COVID-19_publications_datasets_and_clinical_trails/11961063 (2020)
3. Wu Z, McGoogaan JM (2020) JAMA 323:1239–1242
4. Claassen J et al (2019) N Engl J Med 380:2497–2505
5. Sitt JD et al (2014) Brain 137:2258–2270
6. Peiffer-Smadja N et al (2019). Clin Microbiol Ingect. https://doi.org/10.1016/j.cmi.2019.09.009
7. Belhadi D et al (2020) Preprint at https://doi.org/10.1101/2020.03.18.20038190
8. Liu X, Wang X (2020) J Genet Genom 47:119–121
9. Computational predictions of protein structures associated with COVID-19. Deepmind https://deepmind.com/research/open-source/computational-predictions-of-protein-structuresassoicated-with-COVID-19 (2020)
10. Peiffer-Smadja N et al (2020). Clin Microbiol Infect. https://doi.org/10.1016/j.cmi.2020.02.006
11. Stokes JM et al (2020) Cell 180:688–702e13
12. Weiskopf D et al (2020) Preprint at https://doi.org/10.1101/2020.04.11.20062349
13. Senior AW et al (2020) Nature 577:706–710
14. Gautret P et al (2020) Int J Antimicrob Agents. https://doi.org/10.1016/j.ijantimicag.2020.105949
15. Ai T et al (2020) Radiology. https://doi.org/10.1148/radio1.2020200642
16. Chen Z et al (2020) Eur J Radiol 126:108972
17. Pham HH, Le TT, Tran DQ, Ngo DT, Nguyen HQ (2019) Preprint at https://arxiv.org/abs/1911.06475
18. Zheng C et al (2020) Preprint at https://doi.org/10.1101/2020.03.12.20027185
19. Chen T, Kornblith S, Norouzi M, Hinton G (2020) Preprint at https://arxiv.org/abs/2002.05709
20. COVID-19 Clinical Research Coalition Lancet 395:1322–1325 (2020)
21. Sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak. Wellcome Trust https://ac.uk/coronavirus-covid-19/open-data (2020)
22. Open-access data and computational resources to address COVID-19. National Institutes of Health https://datascience.nih.gov/covid-19-open-access-resources (2020)

Role of Internet of Things and Machine Learning in Finding the Optimal Path for an Autonomous Mobile Robot

Dadi Ramesh, Mohmmad Sallauddin, Syed Nawaz Pasha, and G. Sunil

Abstract Path planning for the mobile robot is an emerging area in today's world; the development of autonomous vehicles like driverless cars and mobile robots has tremendously enhanced researchers to work more on path planning strategies using emerging technologies, like the Internet of things and machine learning. These technologies will provide optimal solutions than classical problem-solving algorithms. The article deals with path planning for mobile robots in an unknown environment using deep learning and the Internet of things. These technologies are adopted to work in different strategies like environmental prediction, object detection/ obstacle detection, and finding a path. For this, an innovative model is proposed to detect static and dynamic obstacles from the input data and finding a path from one point to another point.

Keywords Internet of things (IoT) · Machine learning (ML) · Deep learning (DL)

1 Introduction

Path planning is a very crucial task for autonomous mobile robots. The robots always try to find a path from source to destination by overcoming the obstacles and finding an optimal route. This mobile robot should concentrate on predicting the environment and constraints to reach the goal. Today, path planning for driverless cars and autonomous robots is a challenging task. Researchers are working on different types of algorithms to find the optimal path. All the path planning algorithms use traditional approaches like a greedy approach and heuristic approaches to some other algorithms. These are working well, but computational time is more and more space to find a path, and there are not using artificial intelligence techniques [1, 2]. In

---

D. Ramesh (✉)
Center for Artificial Intelligence and Deep Learning, Computer Science and Engineering, S R Engineering College, Warangal, India
e-mail: dadiramesh44@gmail.com

M. Sallauddin · S. N. Pasha · G. Sunil
Computer Science and Engineering, S R Engineering College, Warangal, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_29

this article, the proposed system tends to analyze the role of IoT and deep learning techniques to find the best path.

This paper is organized in the following manner; the related works are explained in Sect. 2. Section 3 discusses the proposed system and role of deep learning and IoT in path planning. Results and analysis are depicted in Sect. 4. Finally, Sect. 5 describes the conclusion of the proposed method.

2 Related Work

The path planning algorithms that will use different technologies to find a path which can classify into three categories, namely the traditional approach, cognitive-based, and artificial intelligence-based methods (Table 1).

2.1 Blend Search

It is an uninformed search where it does not have an idea about adjacent nodes blindly; it searches for goal node. And it does not provide the optimal path. It searches for all possibilities to find the goal node. Such algorithms are breadth-first search [3] and depth-first search [4] which blindly search for goal node without knowing the environment. The computational time is also high, and some no guarantee to find the goal node.

2.2 Grid-Based Method

It follows the concept of state-space search also called graph-based search, representing the environment in an $n * n$ grids and starting cells and goal state in that. It starts searching from starting cell to goal cell. The robot can move cell by cell obstacles are also represented in a cell. It is also an uninformed search, and it searches for the goal in all directions to find the optimal path. This method follows the algorithm which solves the graph and does not try to learn the patterns. And the computational

Table 1 Autonomous mobile robot path planning algorithms

Traditional approach	Cognitive-based methods	AI-based methods
Blend search	Cognitive-based adaptive path planning algorithm (CBAPPA)	D* and A* algorithm
Grid-based method		Ant colony optimization
Divide and conquer method		BEE colony optimization and gray wolf optimization

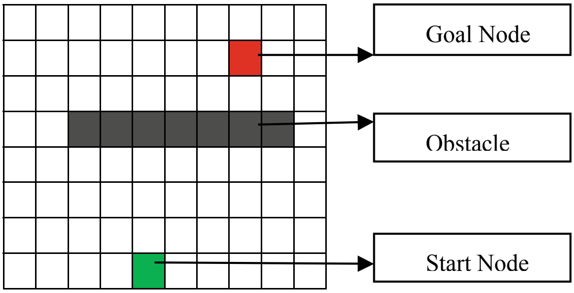


Fig. 1 Environment representation

time is more in finding a path. Figure 1 illustrates the environmental representation of a mobile robot.

2.3 *Divide and Conquer Method*

The divide and conquer method is a popular method to solve the complex problems. The main aim of the divide and conquer [5] way is dividing the problem into a small number of parts and solving the pieces one by one. The same concept was used in mobile robot path planning. It divides the environment into small parts called the local area. First, it finds the path for in local area; later, it merges all local paths to find the final path.

2.4 *Cbappa*

Cognitive-based adaptive path planning algorithm [6, 7] is an algorithm that finds a path from source to destination based on cognitive methods like how a human being perceives the target when he/she does not know the environment. It will choose one direction (goal direction) [8] and starts moving to reach the destination by overcoming the obstacles. It does not search for all possible paths to find an optimal path. It also searches more environments to find the path. But it works in both environments like known and unknown environments.

2.5 *D* and A* Algorithms*

A* [9] and D* [10] algorithms are heuristic search algorithms; A* works for reducing the length of the path by considering all possible paths. And D* practices on reducing the cost of the path from source to destination. A* works on heuristic function $h(x)$. Both are finding optimal paths but searching for more areas to find the final path.

2.6 Ant Colony Optimization

Marco Dorigo proposes this concept. ACO [11] is used to solving computational problems like robot path planning and network path planning. ACO system works on chemical essence, which is released by ants. This phenomenon works how an ant colony finds the best path from their original place to the goal state. The first ant colony starts moving in all directions by releasing pheromones until reaching the goal state. Then, it will decide the optimal path; in the second iteration, ant colony will move in the best path. So ACO's first ants are searching for more areas to find the best path. Almost ants search 60–70% other areas from the actual path to find the optimal path.

2.7 Bee Colony Optimization

The bee colony optimization [12] algorithm proposed by D. Karaboga is based on how the bees find an optimal path from source to destination. The bee colony algorithm works the local search concept first. It searches in the local area later; it increases the local area step by step until reaching the goal state.

2.8 Gray Wolf Optimization

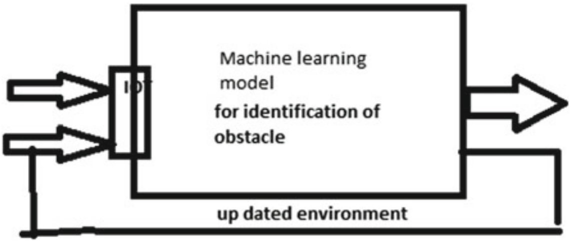
Gray wolf optimization [13] algorithm is based on natural human life how a human being will solve the problems in real time. It works based on a chain pattern. But these two algorithms are finding an optimal path by searching more areas. BCO and GWO algorithms are exploring almost 60% more area than the actual path.

The traditional algorithms can find the path from source to destination, but the complexity is high and not optimal. The AI-based algorithms are finding an optimal way, but the searching area is high. And the CBAPPA algorithm finds the optimal path in some cases. All paths finding algorithms' efficiency are almost less than 50% though they are finding an optimal route. An algorithm's ability is calculated as the area searched for the final path and length of the closing route.

3 Proposed Approach

An IoT and machine learning-based system are proposed for finding a path from source to destination. The selection is based on one path, which is toward the goal point; this path is the primary path. A refined path is identified from a primary path

Fig. 2 System architecture



by overcoming the obstacle and updating the environment. Figure 2 illustrates the autonomous agent architecture with sensors and effectors.

3.1 Obstacle Size Prediction

The input and output modules are controlled by the Arduino microcontroller [14, 15]. It will take the input in a video format and make them into frames. These frames are given to deep neural networks for finding the output path. When the model detects an obstacle, it tries to predict the obstacle size with the help of the prediction algorithm. The algorithm takes the input of three parameters such that the parameter denotes the angle from the local path, the second parameter denotes the distance traveled from a local path, and the third parameter denotes the required angle to reach the local path. In path finding systems, the obstacle size prediction [16] is an important part that will reduce the searching area by predicting obstacle size early. Figure 3 illustrates the methodology of obstacle prediction with the side angle method.

Here, in Fig. 3:

b—is obstacle,

A—is an angle made by a robot corresponding to the primary path.

c—is the distance traveled by a robot from the primary path,

Predicting obstacle size from the image:

$a = \mathrm{BC}$, $b = \mathrm{AC}$, $c = \mathrm{AB}$.

$C = 180 - (\mathrm{A} + \mathrm{B})$

$$b = \frac{asinB}{\sin A}$$

$$\frac{a}{sinA} = \frac{c}{sinC}$$

$$a = \frac{csinA}{sinC}$$

$$\Delta ABC = \frac{1}{2}\mathrm{ac}\sin B$$

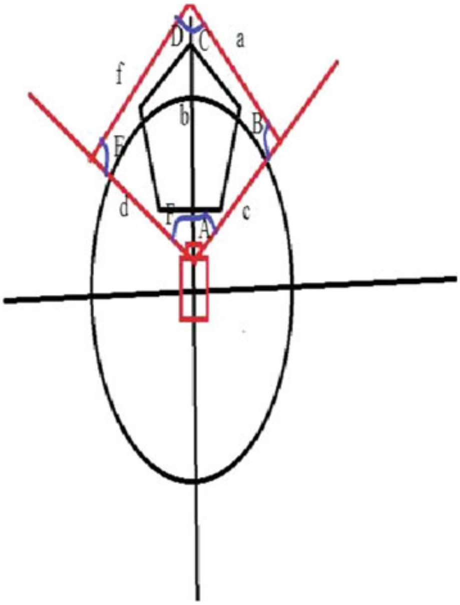


Fig. 3 Obstacle size prediction

From Fig. 3 angles $B \equiv E$.
And $b = \mathrm{DF}, f = \mathrm{DE}, d = \mathrm{EF}$.

$$f = \frac{b \sin F}{\sin E}$$

$$\Delta DEF = \frac{1}{2} \mathrm{bf} \sin D$$

Obstacle size $= \Delta ABC + \Delta DEF$.

3.2 Deep Neural Network Model

The OpenCV and TensorFlow libraries are utilized for implementing. With OpenCV, the video is converted into several frames [17, 18]; after that, each frame is converted to RGB or grayscale image type. Then, it will pass to the input layer.

3.2.1 Training

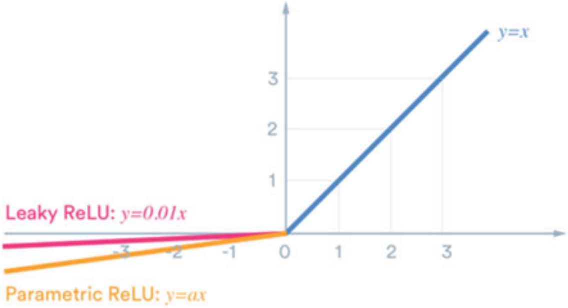
The AlexNet pre-trained model is used to train the images with five convolution layers [19] with max-pooling and zero padding. And the activation function is Relu.

Based on the robot moving, the images will be passed to the model to select the right direction. If the robot moves on a primary path, then the model will find the only obstacle on the primary path until it finds any impediment on the primary path. If it finds any barriers on the primary path, it selects the images from the left and right to move and change the direction.

The trained AlexNet [20] for 12 epochs with dropout layers at each time, the nodes in the convolution layer will be dropped to train the model effectively. The AlexNet model is trained on the dataset, obstacles, free, left direction, and right direction images. Whenever it finds the obstacle first, it calculates the obstacle size and then moves either left or right direction. Whenever it finds the obstacle, it will choose the right or left direction concerning the primary path. And it does not consider the other path.

The test is not executed directly in the robot to find a path from source to destination. Figures 4 and 5 illustrates Relu activation function and CNN architecture corresponding.

Fig. 4 Relu activation function



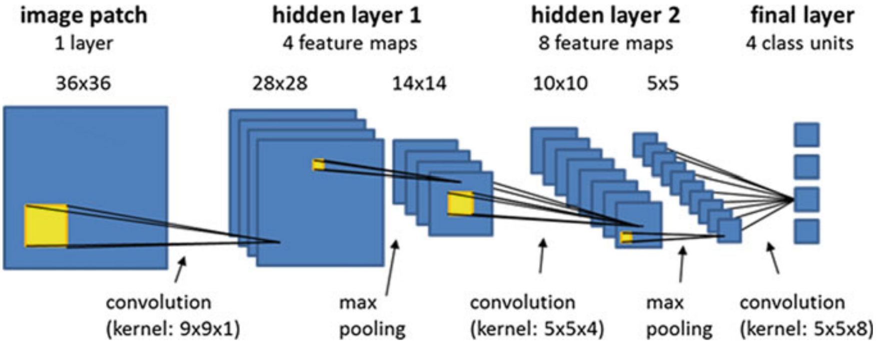


Fig. 5 Sample CNN model

4 Result Analysis

The proposed method is concentrated only on reducing the searching area for finding the final path. The model is trained in a way that whenever the ambiguity occurred, choose to move left or right, the robot will move based on the size of the obstacle. With this, the robot's efficiency will be increased, and the searching area will be reduced by up to 20%.

$$\text{searched area} = \frac{\text{no.of cells searched}}{\text{total no.of cells}} \times 100$$

$$\text{efficiency} = \frac{\text{length of the finalpath}}{\text{number cells searched}} \times 100$$

5 Conclusion

The proposed method presents an approach with Internet of things and a deep learning-based model. It works on unknown environments and finds an optimal path from source to destination, by reducing the searching area. The size of the obstacles is predicted by using trigonometric concepts. When the obstacle size is observed, then the robot will feel like it is moving on known environment, and it is easy to reduce the searching area. It just searches 20% of the new area than the final path.

References

1. Al SNP et al (2019) Variation analysis of artificial intelligence, machine learning and advantages of deep architectures. Int J Adv Sci Technol 28(17):488–489
2. Yue P et al (2019) Experimental research on deep reinforcement learning in autonomous navigation of mobile robot. In: 2019 14th IEEE conference on industrial electronics and applications (ICIEA), IEEE, pp 1612–1616. DOI.org (Crossref). https://doi.org/10.1109/ICIEA.2019.8833968
3. Hansen EA, Zhou R (2007) Anytime heuristic search. J Artif Intell Res (JAIR) 28:267–297
4. Al SM et al (2019) A comprehensive study on traditional Ai and Ann architecture. Int J Adv Sci Technol 28(17):479–487
5. Stentz A (1996) Optimal and efficient path planning for partially-known environments. In: Proceedings IEEE international conference on robotics and automation, pp 3310–3317
6. Ramesh D, Pasha SN, Sallauddin MD (2018) Cognitive based adaptive path planning for mobile robot in dynamic environment. In: Artificial intelligence and cognitive computing
7. Pasha SN, Ramesh D, Roopa G A novel approach to path planning of robots by electing dynamic obstacles
8. Manoharan S, Ponraj N (2019) Precision improvement and delay reduction in surgical tele robotics. J Artif Intell 1(01):28–36

9. Goldberg AV, Harrelson C (2005) Computing the shortest path: A* search meets graphtheory. In: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, pp 156–165
10. LaValle SM (2006) Planning algorithms. Cambridge University Press, Cambridge, UK; Stentz A (1995) The focussed D* algorithm for real-time replanning. Int J Rob Autom
11. Gambardella LM, Dorigo M (1997) Ant colony system: a cooperative learning approach to the traveling salesman problem. IEEE Trans Evol Comput
12. Teodorović D (2009) Bee Colony optimization (BCO). In: Lim CP, Jain LC, Dehuri S (eds) Innovations in swarm intelligence. Studies in computational intelligence, vol 248. Springer, Berlin, Heidelberg
13. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. J Adv Eng Softw 69:46–61
14. Li Q, Chen L, Li M, Shaw SL, Nüchter A (2013) A sensor-fusion drivable-region and lanedetection system for autonomous vehicle navigation in challenging road scenarios. IEEE Trans Veh Technol 63(2):540–555
15. Smys S, Ranganathan G (2019) Robot assisted sensing, control and manufacture in automobile industry. J ISMAC 1(03):180–187
16. Adapting Best Path for Mobile Robot By Predicting Obstacle Size (2019). Int J Innov Technol Explor Eng 8(9S2):200–202. DOI.org (Crossref). Blue Eyes Intelligence Engineering & Sciences Publication. https://doi.org/10.35940/ijitee.I1039.0789S219.
17. Fujiyoshi H et al (2019) Deep learning-based image recognition for autonomous driving. IATSS Res 43(4):244–252. DOI.org (Crossref). https://doi.org/10.1016/j.iatssr.2019.11.008
18. Harshavardhan A et al (2020) 3D surface measurement through easy-snap phase shift fringe projtion. Springerprofessional.De. https://www.springerprofessional.de/en/3dsurfacemeasurement-through-easy-snap-phase-shift-fringe-proj/15447362. Accessed 26 Mar 2020
19. Levine S, Finn C, Darrell T et al (2016) End-to-end training of deep visuomotor policies. J Mach Learn Res 17(39):1–40
20. Ren S, He K, Girshick R, Sun J (2017) Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 39(6):1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031.

Overview of Blockchain Technology: Applications and Use Cases

J. S. Shyam Mohan, Vedantham Hanumath Sreeman, Vanam Venkata Chakradhar, Harsha Surya Abhishek Kota, Nagendra Panini Challa, M. U. M. Subramanyam, and Surekuchi Satya Swaroop

Abstract Blockchain technology has revolutionized agriculture, finance, education, supply chain, health care, and many sectors. Blockchain has extended its advantages to the non-financial sector too. Researchers in this area are constantly trying to explore many applications of blockchain that can provide fruitful results. This article discusses various use cases and real-time applications of blockchain technology in various sectors and subsequently, explores the state-of-the-art developments in blockchain use cases in various sectors and the scope for further research aspects in the future. It is intended for audience those who are interested to learn use cases and real-time applications of blockchain technology.

Keywords Distributed ledger technology (DLT) · Federated Byzantine Agreement (FBA) · Monetary Authority of Singapore (MAS)

J. S. Shyam Mohan · V. H. Sreeman · V. V. Chakradhar (✉) · H. S. A. Kota ·
M. U. M. Subramanyam · S. S. Swaroop
Department of CSE, SCSVMV, Kanchipuram, India
e-mail: vanamvenkatachakradhar@gmail.com

J. S. Shyam Mohan
e-mail: jsshyammohan@kanchiuniv.ac.in

V. H. Sreeman
e-mail: hanumathsreeman@gmail.com

H. S. A. Kota
e-mail: khsabhishek1335@gmail.com

M. U. M. Subramanyam
e-mail: mani12061999@gmail.com

S. S. Swaroop
e-mail: satyaswaroop058@gmail.com

N. P. Challa
Department of IT, SVECW, Kovvada, India
e-mail: paninichalla123@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_30

1 Introduction

With the tremendous improvements in blockchain technology and with the increasing utilization of bitcoins and continuous improvement in blockchain technology, various academic and financial sectors are continuously exploring the practical applications of blockchains. Today, the majority of the financial, sales, and clinical fields' blockchain has shown its impact. In this paper, the authors focus on the applications and use cases of blockchain technology in various sectors. Transactions in blockchain are handled in the distributed ledgers using tokens. Apart from using tokens, Ethereum and Hyperledger Fabric are also used for making digital transactions [1, 2]. Blockchain has been commercially adopted, influenced the world's currency market, and facilitated the illegal dark web. Blockchain additionally has been a critical factor influencing the expansion of monetarily determined digital assaults, for example, ransomware and denial of service against retailers and other online associations. The usage of blockchain has brought the world's first decentralized cryptocurrency. Mostly, this is because of hype, powered by the rising and dropping estimation of bitcoin. Blockchain applications play a vital role in 5G and 6G technologies. The impacts of distributed technology on business systems give off an impression of being like dis-intermediation and digital intervention impacts in e-commerce (Laudon and Traver 2018). The blockchain applications may have an assortment of consequences for business systems and business relations, remembering impacts for trust and consequences for the system structure.

Research in distinctive business regions and use cases where blockchain technology leads to the transformation of a new era. The implementation of blockchain technology in various fields will be helpful for advancement and a trustworthy environment. Many applications in blockchain are based on distributed ledger technology (DLT). Blockchain can increment money-related productivity by diminishing manual control and tampering of the existing data. In intercompany trades, blockchain will make one type of the record allowing intercompany straightforwardness and settlement at a comparative second. Blockchain is an open record that keeps up records of the impressive number of trades held tight a blockchain mastermind while working in a scattered manner. This blockchain coordinator is a conveyed framework that does not waste time with any central capacity to confirm or settle the trades in the crucial framework and thusly empty go-betweens and bring straightforwardness and improved security. Blockchain technology has been adopted for use in many government sectors. Because of digitalization, character the board has consistently involved worry for all open just as private associations. Blockchain applications and use cases by industry are shown in Table 1.

This paper is organized into the following sections: Sect. 2 describes the blockchain technology and use cases for payments in financial services, and Sect. 3 records the compliance and mortgage. Section 4 represents the blockchain for Global Trade Logistics. Section 5 outlines the blockchain in Healthcare, and Sect. 6 explains the blockchain in energy markets. Section 7 depicts the blockchain in government. Finally, Sect. 8 concludes the research work.

Table 1 Blockchain applications and use cases by industry

Financial services	Insurance	Retail	Supply chain and logistics	Public sector
Trade finance	Claims processing	Supply chain	Supply chain and finance	Asset registration
Cross-currency payments	Risk provenance	Loyalty programs	Maintenance tracking	Citizen identity
Mortgages	Asset usage history	Information sharing	Provenance	Medical records
KYC	Claims file		Supply chain compliance	Medicine supply chain
Cross-border taxes				

2 Blockchain Use Cases—Payments and Securities Trading in Financial Services

Blockchain technology has been applied to banking and money-related administrations in different manners and getting various advantages. Keen agreement administration helps in leading budgetary exchanges without a go-between. It can oversee protections, deeds, settlements, and cases in a robotized way. With the advanced development in the financial area, blockchain innovation models have just been executed in worldwide installments which are profiting banks as far as diminishing expenses and to abbreviate handling times. The economy purely depends on the circulation of the respective country currency, and the overview of financial services that inherited blockchain technology is as follows [3].

2.1 Cross-Border Payments—Stellar Open Network

A decentralized, hybrid blockchain platform that makes it possible to create, send, and trade digital settlements on a single network with Lumens as a native asset is based on Federated Byzantine Agreement (FBA) that takes approximately 2–5 s for each transaction clearance. It has anchors that act as bridges between a given currency and a Stellar network that consists of a distributed exchange, viz., pay in EUR with INR balance, and the network will automatically convert it at the lowest available rate.

Figure 1 shows the Stellar performance that it could be the best solution and might this blockchain technology can replace the interbank transfer method which was vulnerable in some cases, and with the cryptocurrency security, these types of frauds could be nullified. The Stellar network stock fluctuations represent the currency circulations and depend on the particular country currency value at the international

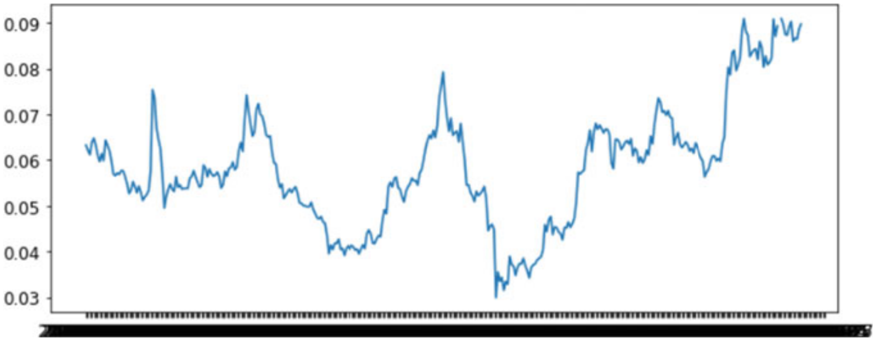


Fig. 1 Graph of Stellar EUR performance

level and the conversions made by the users, and these all factors involved in the performance of the Stellar.

2.2 Permissioned Network for Payments

Maintains Nostro (“Ours”) and Vostro (“Your”) accounts and transactions. Only member banks are permitted to transact, ensuring privacy and confidentiality of transactions, and eliminate the need for reconciliation and errors that happen, reducing costs and delays in clearing and settlement (like RTGS). The treasury has an instantaneous view of the currency position of their Nostro accounts across the globe, allowing for optimal use of capital. It provides reduced foreign exchange and capital exposure, lower fees, and increased compliance and security [4].

2.2.1 Project Ubin: Central Bank Digital Money Using Distributed Ledger Technology

A synergistic venture with the business to explore the utilization of blockchain and distributed ledger technology (DLT) for clearing and settlement of installments and protections. Project Ubin aims to assess the ramifications of having a tokenized type of the Singapore Dollar (SGD) on a distributed ledger (DL) and its expected advantages to Singapore’s monetary biological system. Project Ubin was considered as an open door for Singapore to play the main job in the exploration of national bank money on a DL and Central Bank Digital Currencies (CBDCs). Monetary Authority of Singapore (MAS) is Singapore’s central bank. MAS goes about as a settlement specialist, administrator, and manager of an installment, clearing, and settlement frameworks in Singapore that attention on wellbeing and effectiveness. MAS embraces this job as a confided in an outsider (trusted third party) and effectively connects with banks in Singapore, just as with the open and private sector, for example, the Singapore Clearing House Association (SCHA) and the Association of Banks in Singapore (ABS). Since MAS has become a more trusting third party, this introduced an incredible open door for MAS to work together with the banks and survey the worthiest that blockchain could bring to this current relationship [5]. Some of the phases of Project Ubin are Tokenized are SGD, Re-imagining RTGS, Delivery versus Payment (DvP), Cross-border Payment versus Payment (PvP), Enabling broad ecosystem collaboration [6].

2.3 *Ripple Protocol and Network—Financial Services*

Ripple protocol is used for banks to clear and settle payments in real time through a distributed network without a centralized clearance house. It takes an average of 5 s for confirmation. Gateway nodes convert fiat currencies to XRP (currency in Ripple) (Fig. 2).

Interbank payment depends on the daily fluctuations in the currency of a country in international trades and makes a linked path which confuses the process and the centralized system monitoring the transactions of the investors/organizations. The security of the transactions is not compromised, but the personal information and transaction details of an individual investor are recorded, and there might be a chance of losing investors due to monitoring their records. The traditional method does have some time-taking process because every transaction made by the person or investor

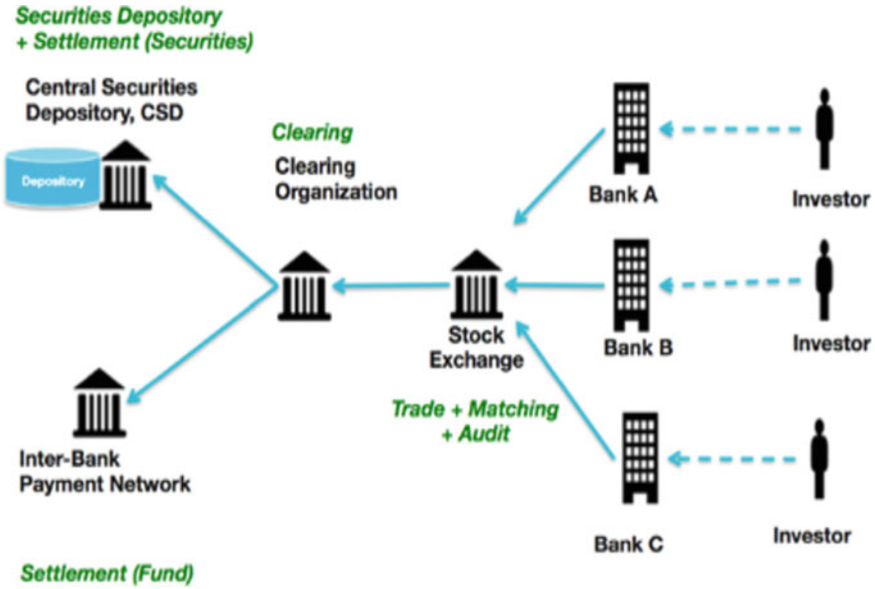


Fig. 2 Securities settlement—today's process [7]

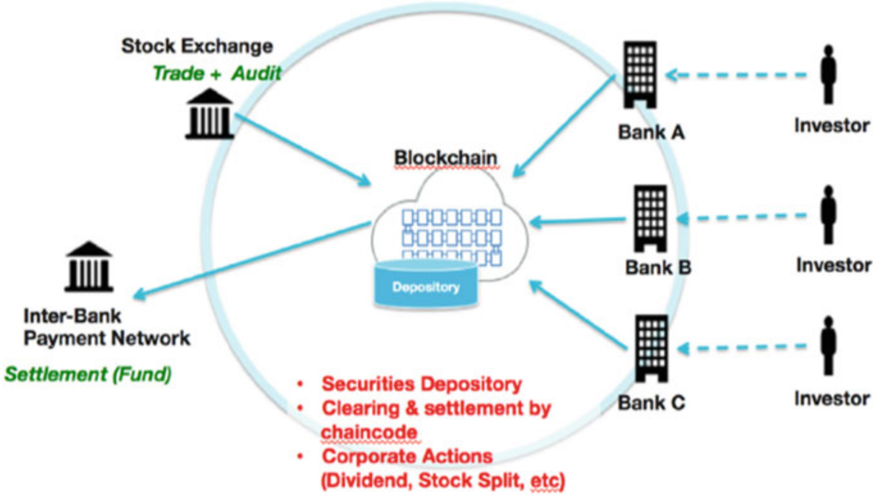


Fig. 3 Securities settlement—blockchain process

is centralized and because many people involved brags the transaction time more [7] (Fig. 3).

The blockchain process provides a more secure way of digital transactions with all the cryptographic records, and every investor's transaction details are secured in between them, and no one is there to monitor the transactions. The inter-border fund transactions would be easier and secure than the complex traditional way, for example, Stellar can precede the transaction conversions according to the currency trades between the persons involved in the transaction. This method also provides a method of unified currency options like bitcoins which then converted into their corresponding country currency. Blockchain is a type of distributed ledger, distributed ledger method, which provides the cryptographic documents/transactions between the two persons or investors or dealers. There is no central administrator in a distributed ledger, mainly focused on digital transactions, records, and security. Distributed ledgers are like databases that are synchronized, duplicated, and shared; there are no others to change and record the data between the shared persons [7].

3 Compliance and Mortgage

A Mortgage is a kind of agreement that tends if the client neglects to reimburse the credit sum with interest taken from the associations, and then, the associations have the right to take the property referenced as an affirmation in the advancement procedure. Blockchain can be utilized to make a computerized ID for every property, in this way making the property trackable on the system. Aside from making the land showcase more fluid, simply from a home loan application point of view, this computerized ID would incorporate a chain of possession and a current market valuation that will permit banks to rapidly check the current proprietorship status or affirm the market cost, conceivably relieving the need of experiencing title deeds and drawing in with assessors.

3.1 Compliance (KYC and AML)

Know Your Customer is a process of verifying the due diligence of a client. Sharing KYCs on a blockchain would provide financial institutions to deliver better compliance outcomes and improves customer experience [8, 9]. KYC compliance is of the following steps they are gathering of information, verification and due diligence, initial risk assessment, and continuous monitoring (Fig. 4).

Blockchain innovation takes into account the making of a distributed ledger that is then mutual with all clients on the system. Ethereum is a blockchain stage that utilizes a smart contract for handling every single exchange. This factor implies that nobody has single power, and consequently, a state of shortcoming, as in the client/server model, and also the information are immutable. This implies blockchain databases have an inbuilt changelessly that makes the information that they contain unquestionably more dependable. Such databases can be utilized to store ID subtleties of people which would be dependable. On the off chance that the money-related administration area, for instance, executes blockchain for KYC confirmation, they will have the option to check clients rapidly and dependably by using an application, and so on. Because of the unwavering quality of blockchain databases, government establishments and organizations could depend on the information totally, something which would evacuate the requirement for any further ID checks [10].

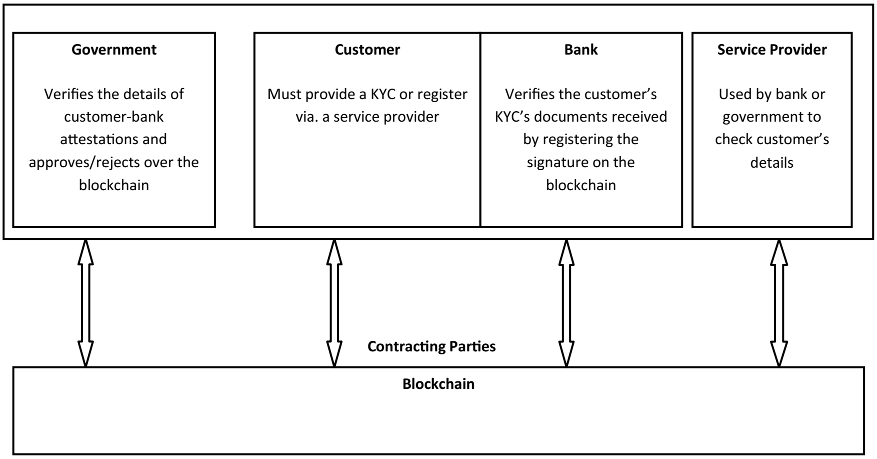

Government
Verifies the details of customer-bank attestations and approves/rejects over the blockchain
Customer
Must provide a KYC or register via. a service provider
Bank
Verifies the customer's KYC's documents received by registering the signature on the blockchain
Service Provider
Used by bank or government to check customer's details
Contracting Parties
Blockchain

Fig. 4 Compliance (KYC and AML)

3.2 Trade Finance Network

The platform is built on the IBM Blockchain Platform using Hyperledger Fabric that provides easy access to banking services to customers. It takes into consideration adaptability that takes into consideration fast universal extension as a business, administrative, and security opportunities converge [11, 12].

4 Blockchain for Global Trade Logistics

To work internationally, the shipments are required to proceed onward time. Blockchain conveys the continuous following of the considerable number of members in the flexible fasten to see the exactness, adjust verification data, and information. DLT helps in making the best stage for the administration of every single part of delivery coordinations. Since the data is immutable, there is no possibility of data tampering. Data in a blockchain cannot be hacked or forged and is promptly "trusted" and along these lines, acknowledged by anybody with access to your chain [13].

4.1 IBM Blockchain for Trade Logistics

IBM Blockchain for trade logistics provides an effective way of transporting goods from one location to another location [14]. Some of the services provided by IBM Blockchain are container logistics, food supply, procurement, responsible sourcing, counterfeit prevention, and supply chain visibility.

4.2 Tradelens

Tradelens is powered by IBM Cloud and IBM Blockchain. It gives each substance associated with a worldwide exchange with the computerized instruments to share data and team up safely.

5 Blockchain: Healthcare

Distributed ledger technology is used to store the information in an immutable format and updates the data progressively which has been reshaping the human services area entirely. The traditional models in this scene end up being exceptionally wasteful as far as conveying quality human services which are moderate in nature of the people. Blockchain technology-based human services applications are fit to be utilized and change the social insurance establishments over the world Blockchain technology works for the improvement as far as straightforward and proficiency; different gatherings are related to the medicinal services framework, and patients get profited [15].

5.1 GuardTime

Guardtime HSX crosses over any barrier between patients, suppliers, payers, controllers, and pharma via flawlessly shipping information over various human services partners, conveying secure utilization of a solitary, honest form of wellbeing information [16].

5.2 Loyyal

With Loyyal's Blockchain-as-a-Service (BaaS), customers have comprehensive access to undertake grade facilitating administrations, advancement apparatuses, bolster administrations, and ever-developing system of accomplices. Comprehensive month-to-month permit expenses can begin as low as $5 K every month suitable for particular needs. The following are some of the features of loyal, namely unlimited API access to the loyal platform, entry to loyal's network of earning and redemption partners, personalized node dashboard, monthly support services, and unlimited support for severity level 1–2 issues.

6 Blockchain in Energy Markets: Gridchain

PONTON has developed imaginative pilot programming dependent on blockchain innovation that simulates future forms for constant lattice the executives, called Gridchain. The next stage will be that for all intents and purposes, test this procedure in the field "with a choice of market members, i.e., an alliance of the willing." Also, Gridchain is a commitment to the European institutionalization of between process correspondences when planning keen lattices of things to come [17].

6.1 Bloomberg New Energy Finance (BNEF)

Bloomberg New Energy Finance (BNEF) produces investigation on ventures on the move, concentrating on clean vitality, propelled transport, advanced industry, creative materials and, wares. BNEF bits of knowledge help corporate methodology; money and strategy experts separate the truth from the publicity, explore change, and produce openings. The BNEF controls the accessibility, portability, and the terminal by 250 examiners in 18 areas around the world. Figure 5 shows detailed information about the BNEF [18].

The network edge, where the person connects, and the distribution grid balance the supply and demand with no central controller, and Bloomberg New Energy Finance helps to create opportunity very useful while including in the corporate strategy, finance, and policies. Bloomberg is always interested in fields like clean energy, digital industry, advanced transportation, and innovative products. For example, the bulk generation in Fig. 5 had solar panels, a windmill plant registry with fraud-proof, the carbon tax for the transport, and energy sources collected directly from appliances, and all transmission networks resolve all the payments peer to peer.

7 Blockchain in Government

Blockchain [19] helps the government in the following ways such as access and verification of central data, sharing of data, and sharing of data and access control.

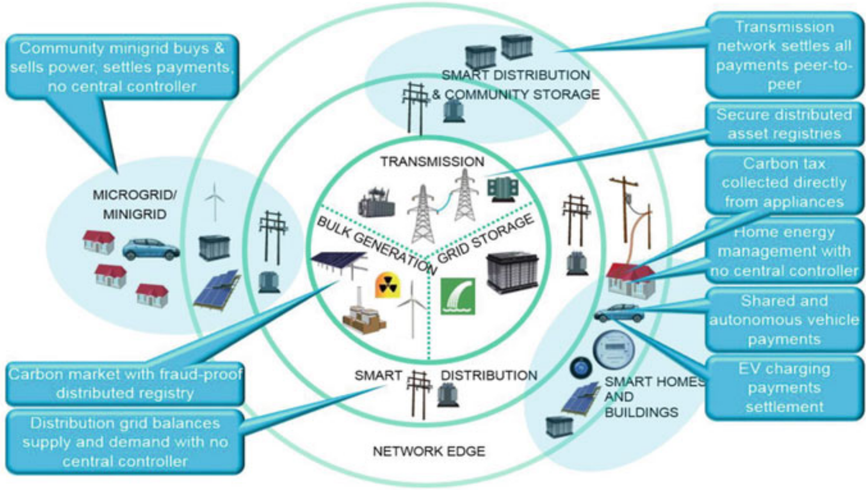

Community minigrid buys & sells power, settles payments, no central controller
SMART DISTRIBUTION & COMMUNITY STORAGE
Transmission network settles all payments peer-to-peer
Secure distributed asset registries
Carbon tax collected directly from appliances
Home energy management with no central controller
Shared and autonomous vehicle payments
EV charging payments settlement
MICROGRID/ MINIGRID
TRANSMISSION
BULK GENERATION
GRID STORAGE
SMART DISTRIBUTION
SMART HOMES AND BUILDINGS
NETWORK EDGE
Carbon market with fraud-proof distributed registry
Distribution grid balances supply and demand with no central controller

Fig. 5 Bloomberg new energy finance

7.1 Blockchain Use Cases in Government—Worldwide

7.1.1 Russia

The first government-level blockchain implementation is completed officially in Russia. The state-run bank Sberbank announced today that it is partnering with Russia's Federal Antimonopoly Service (FAS) to actualize record move and capacity via blockchain.

7.1.2 South Korea

South Korea's leading blockchain project is a network designed to interconnect independent chains and their communities; in essence, ICON is a blockchain of blockchains. Consensus models are continuously developing and one of the most Delegated Proof-of-Contribution (DPoC) which takes delegated proofs of stake above and beyond.

7.1.3 Singapore

The government has initiated a Project Ubin that is built on the distributed ledger for clearing and settlement of installments and protections. DBS Bank has cooperated with the Singapore government for the blockchain trade platform.

7.1.4 India

IndiaChain is the first blockchain initiative proposed by NITI Aayog in India. It is used to maintain India's public records. This has collaborated with UPI for transactions. Andhra Pradesh has entered into a partnership with Chromaway to use blockchain technology to maintain land registry records [20].

7.1.5 USA

Federal agencies in the USA are trying to assess and receive appropriate record advances that utilizes encryption and coding to improve straightforwardness, proficiency, and trust in data sharing like financial management, trademarks, and many more services.

Table 2 Blockchain applications and use case for other sectors

S. No.	Blockchain application	Description
1	Hyperledger Indy	Distributed ledger platform for decentralized identity management
2	Hyperledger Indy—Plenum	Distributed ledger platform that uses Redundant Byzantine Fault Tolerant algorithm for consensus
3	SecureKey	SecureKey is a driving personality and confirmation supplier that rearranges shopper access to online administrations and applications
4	Sovrin	Personally manage individual IDs online
5	IBM Blockchain	Helps to create, operate, and maintain permitted decentralized identity networks built using Hyperledger Indy DLT

7.1.6 Estonia

e-Estonia is like a digital Id card and a decentralized distributed system [21]. Some of the advantages of e-Estonia are file taxes within 5 min, sign a contract electronically, register a business within 30 min, I-voting, and registering as a citizen (Table 2).

8 Conclusion

An overview of blockchain technology applications and use cases in various sectors and the study of Stellar's data shows a better understanding of blockchain technology because some countries like India banned blockchain technology in some fields. Blockchain is a kind of distributed ledger, which gives the cryptographic documents/transactions between the two people or speculators or dealers. Many applications in blockchain depend on distributed ledger technology (DLT). With the serious advancement in the budgetary territory, blockchain development models have quite recently been executed in overall portions which are benefitting banks similar to reducing costs and to abbreviate handling times. However, this article provides insights into modern-day applications of blockchain technology in real time. The interested audience can refer to the references and web links for more details.

References

1. Aste T, Tasca P, Di Matteo T (2017) Blockchain technologies: the foreseeable impact on society and industry. Computer 50(9):18–28
2. Zheng Z, Xie S, Dai H, Chen X, Wang H (2017) An overview of blockchain technology: architecture, consensus, and future trends. In: 2017 IEEE international congress on big data (BigData Congress), p 557564

3. Laroiya C, Saxena D, Komalavalli C (2020) Applications of blockchain technology. In: Handbook of research on blockchain technology, pp 213–243
4. Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system. www.Bitcoin.Org. https://bitcoin.org/bitcoin.pdf [Online]. Available
5. Project Ubin—https://www.mas.gov.sg/-/media/MAS/ProjectUbin/Project-Ubin--SGD-on-Distributed-Ledger.pdf
6. Project Ubin—https://www.mas.gov.sg/schemes-and-initiatives/Project-Ubin
7. Blockchain charts—https://www.blockchain.com/en/charts
8. JPX—https://www.jpx.co.jp/english/
9. Blockchain Applications in Supply chain—https://www.ibm.com/blockchain/industries/supply-chain
10. Sinha P, Kaul A (2018) Decentralized KYC system. Int Res J Eng Technol (IRJET)
11. Loyal—https://loyyal.com/
12. Skuchain—https://www.skuchain.com/about/
13. David MD et al (2018) Blockchain for and in logistics: what to adopt and where to start. Multidisciplinary Digital Publishing
14. Kan L, Wei Y, Hafiz Muhammad A, Siyuan W, Linchao G, Kai H (2018) A multiple blockchains architecture on inter-blockchain communication. In: 2018 IEEE international conference on software quality, reliability and security companion (QRS-C), pp 139145
15. Zeadally S, Siddiqui F, Baig Z, Ibrahim A (2019) Smart healthcare: challenges and potential solutions using internet of things (IoT) and big data analytics. PSU A Rev J 1–17
16. Guard time—https://guardtime.com/health
17. Gridchain—https://enerchain.ponton.de/index.php/16-gridchain-blockchain-based-process-integration-for-the-smart-grids-of-the-future
18. Bloomberg report—https://www.bloomberg.com/impact/products/bloombergnef
19. Blockchain for government—https://digital.gov/communities/blockchain/
20. Indiachain—https://www.businessinsider.in/what-is-indiachain-a-blockchain-system-that-could-soon-be-the-heart-of-governance-in-india/articleshow/64676670.cms
21. E-estonia—https://e-estonia.com/

Tuberculosis Detection from CXR: An Approach Using Transfer Learning with Various CNN Architectures

P. Anu Priya and E. R. Vimina

Abstract Tuberculosis (TB), a major public health threat, is preventable and curable if identified at its earlier stage. Advances in artificial intelligence lead convolution neural network (CNN) to focus on TB elimination by using popular diagnostic tools like chest X-rays. However, the limited availability of publicly accessible chest X-ray datasets remains a challenge that can be resolved using the transfer learning technique. In this work, to detect TB, transfer learning is used with different CNN architectures such as VGG-19, RestNet50, DenseNet121, and InceptionV3 on Montgomery, Shenzhen, and combined dataset. For performance evaluation, the area under the ROC curve (AUC) and accuracy (ACC) along with the confusion matrix is considered. The results show that the VGG-19 model achieved the highest AUC score of 0.89, 0.95, and 0.95 for Montgomery, Shenzhen, and combined datasets, respectively.

Keywords Tuberculosis · Deep learning · Convolutional neural network · Artificial intelligence · Transfer learning

1 Introduction

Artificial intelligence (AI) is becoming the future of mankind, especially in the field of medical science. The standardization of diagnostic measures in disease identification to be specific is always a developing and promising ground for technological advancements like AI. But in the world of evidence-based medicine, all the possibilities need to be tested to its core. At the same time, potentials of this technology should be focused on the eradication of major public health threats like tuberculosis.

Tuberculosis (TB) has always been a persistent public health threat caused by Mycobacterium tuberculosis. Being an airborne disease, TB is transmitted from person to person through cough, sneeze, and spits of infected individuals. The

---

P. Anu Priya (✉) · E. R. Vimina
Amrita School of Arts and Sciences, Amrita Viswa Vidyapeetham, Kochi Campus, Kochi, India
e-mail: priyaanu29@gmail.com

E. R. Vimina
e-mail: vimina.er@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_31

disease is said to be under control in developed countries, but the global condition stays perilous as the developing and underdeveloping countries are still struggling to manage the malady. This statement can be indicated by World Health Organization (WHO) global report statistics which is getting updated every year that says the disease affects nearly 10 million of the world population and kills an estimate of 1.6 million annually [1]. The major contributor for the global burden for TB in India as the country's caseload when added up will make up to at least one-third of the global TB cases. Figure 1 shows the year-wise TB incidents in India from 2010 to 2018. In 2016, 28 lakh Indian citizens became ill with mycobacterium tuberculin and 4.5 lakh death due to the disease occurred [2]. WHO Report from 2019 revealed that about 2.6 million cases occurred in the country, and 0.45 million death occurred due to the disease.

A major cause of high mortality and spread is due to the persistent gap occurring in the detection of this disease. Therefore, it is a known fact that more than one-third of the estimated 10 million cases are not undergone timely diagnosis or reporting. The widely used two diagnosing measures for this disease are Ziehl–Neelsen (ZN) staining test and chest X-ray. Though CXR is having low specificity, it is still a popular measure used for pulmonary TB detection by resource-limited settings of primary-level clinicians and mass screening programs worldwide. Chest X-ray is said to make about 37% over diagnosis when no other diagnostic measures are used along with it to confirm the disease occurrence [3–5]. These can be caused due to factors like unavailability of experts to read the X-ray film, compromised quality of the machine or film used, the gender of the patients, etc. Inter-reader and intra-reader variations in CXR reading are also a notable factor when it calls out for the need of involving artificial intelligence in the disease identification [6, 7]. The current lack

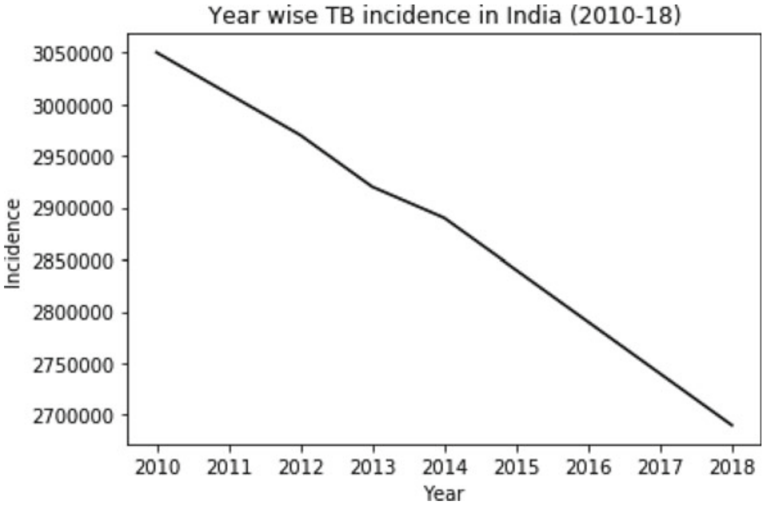


Fig. 1 Year-wise TB incidents in India

of evidence is preventing authority bodies like the World Health Organization from developing recommendations regarding the use of automated reading systems for TB detection [8].

Though deep learning was proving its ground in several studies for the detection of cancer cells and diabetic retinopathy, lack of studies based on TB detection from CXR using deep learning systems made a demand for its involvement in recent years. The researchers were looking forward to the expectation that there would be less inter-reader variability, provision of radiological services in places where experts were unavailable, and reproducibility of valid results. Now, deep neural networks provide new possibilities for TB detection from chest radiographs. Developments in neural networks to classify images, sound, and text with algorithms that use nodes and layers to learn like that of human cognitive functioning have always contributed to this advancement. This is achieved by recurrent training sessions using different datasets [9, 10].

CNN is a technique commonly used in deep learning that arranges nodes in tiles to build a visual reception area. The layers for training extract significant visual definitions of images from data, while sets of metrics to be prepared using fully connected layers classify the characteristics to target groups, e.g., TB or normal [11]. As CNN is programmed to identify and extract the most discriminatory features [12, 13]; based on the target objectives from the data itself, it eliminates the requirement for manual feature inputs that rely on domain-specific knowledge. It is practically difficult to provide enough data to train the deep networks due to the lack of publicly available datasets. This issue can be overcome to a level using transfer learning which uses pre-trained CNN models on large datasets; for example, ImageNet has 1.2 million of 1000 class images. In this technique, pre-trained CNN models use initial weights that are prepared with a large-scale dataset to train CNN models later, even with a small database [14].

The objective of this study is to evaluate the efficiency of advanced pre-trained CNN in TB detection from chest X-ray. To accomplish this, X-ray images from publicly available datasets are processed and used to train and test CNNs. Since TB-related samples are small, the favored technique for training deep CNN is transfer learning. Therefore, the transfer learning approach with four different CNN architectures was used to detect the presence of TB and later to compare the results.

2 Methods

2.1 Dataset

The publicly accessible Montgomery dataset and Shenzhen dataset of NIH Tuberculosis Chest X-ray dataset is used. The National Library of Medicine developed Montgomery dataset by associating with the Health and Human Services Department,

Table 1 Dataset summary

	Montgomery	Shenzhen	Combined
Subjects	138	662	800
Without TB	80	326	406
With TB	58	336	394
Male	63	442	505
Female	74	213	287
Other/unknown	1	7	8

Montgomery County, Maryland, USA, and created Shenzhen dataset in collaboration with the Shenzhen No. 3 Hospital of China [15, 16]. Both sets contain normal and abnormal X-rays, the latter containing tuberculosis manifestations. The dataset of Montgomery and dataset of Shenzhen have 138 and 662 subjects, respectively, with the presence and absence of TB. Table 1 provides information regarding the datasets. For taking advantage of the discrepancies in processing between various sets of data, a combination of both datasets has been created. The model is trained on these datasets to acquire robust functionality.

In CXR images, tuberculosis is seen in the form of infiltrates, cavitations, blunted costophrenic angles, consolidations, pleural effusion, opacities, pneumonia, horizontal fissure displacement, widely spread nodules, and in many other radiological forms [17]. Figure 2 shows CXRs of subjects chosen from the dataset. Figure 2 a exhibits CXRs of subjects uninfected with TB and b, c, d shows CXRs of patients infected with TB with different manifestations.

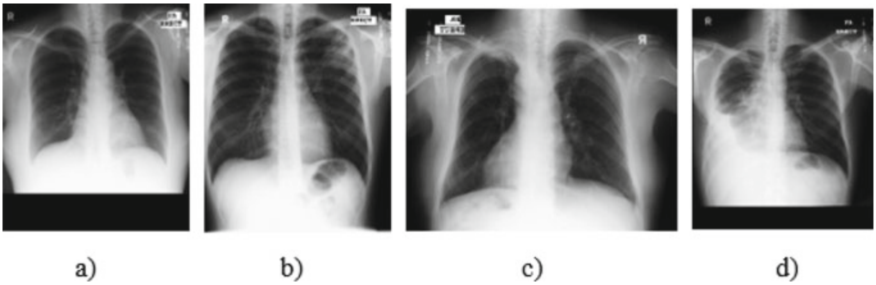


Fig. 2 **a** Normal CXR, **b** cavitary infiltrates, **c** infiltrates, **d** pleural effusion

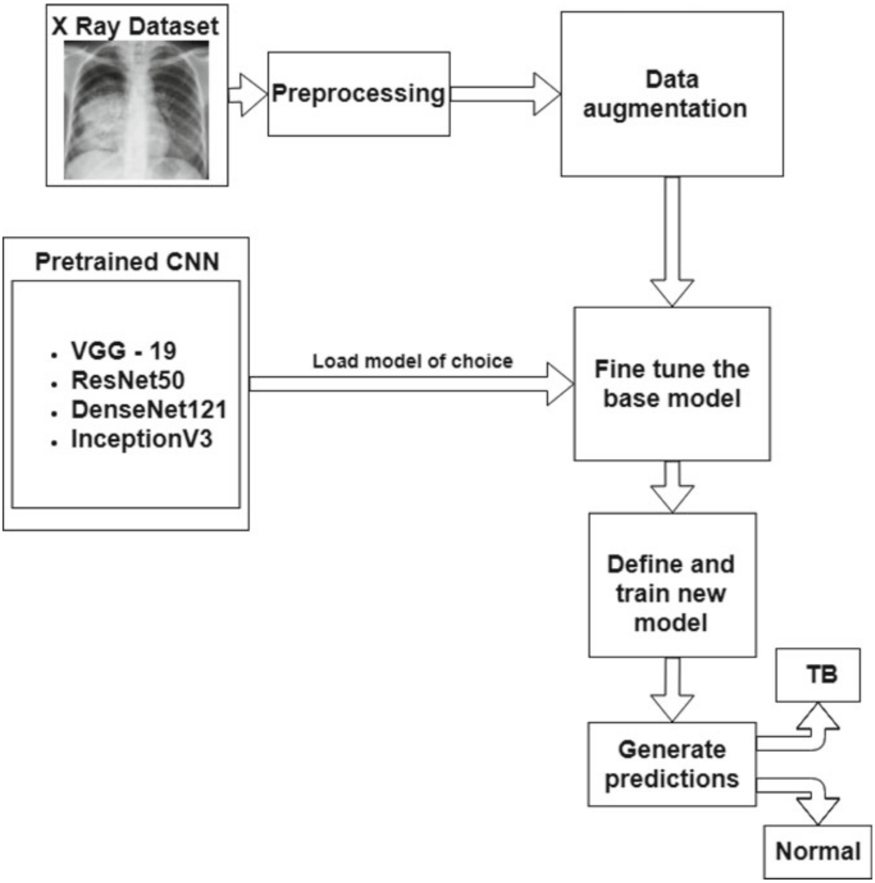


Fig. 3 Conceptual diagram

2.2 *Methodology*

The overall procedure is visualized through a conceptual diagram in Fig. 3. The following steps are included: loading CXR images, preprocessing of CXR, augmentation transformation on training and validation images, transfer learning with VGG-19, DenseNet121, InceptionV3 or ResNet50 architechures, fine-tuning of the base model, defining and training new model, and generate predictions.

2.2.1 Data Preprocessing and Augmentation

All images used in the datasets for this study are frontal thoracic chest X-rays and include regions beyond the lungs that do not apply to TB detection. Preprocessing is performed to diminish the features that are irrelevant to TB detection that distorts the final results. For this purpose, few of the available options were used to initialize data generators for training and validation datasets. In training, data generator transformations are performed on images to produce new images by applying rescaling on images and using the fill mode parameter to fill in new pixels after applying transformations with the nearest surrounding pixel values. All the models pre-trained were very large to carry the dataset of this size, which could easily overfit the model. To prevent this, data augmentation is performed. An iterator is created to train the network model by specifying the target size of the image as 128 * 128, batch size as 64, shuffle to a true value, and class mode set as categorical. No other preprocessing techniques have been carried out.

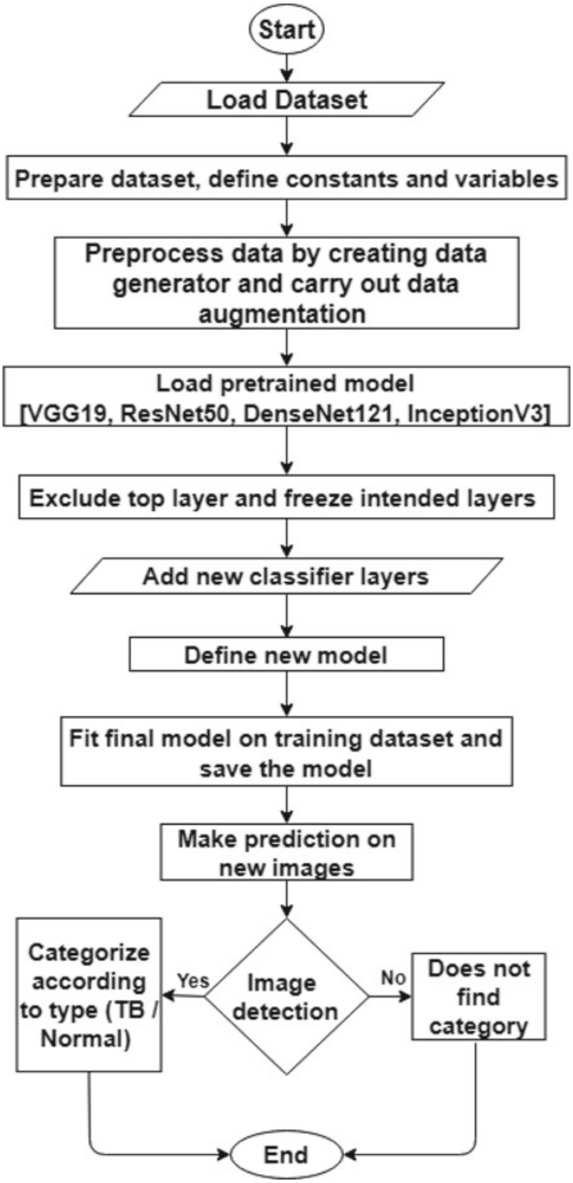
2.2.2 Transfer Learning with CNN

In several computer vision tasks including classification and segmentation, CNN has demonstrated its promising performance. But, since deep learning is a data-driven process, learning the rules involves a great deal of data. If there are not enough training data, it is difficult for deep learning algorithms to optimize the parameters of the prediction models. Working on the related issue of computer vision, the best approach is to involve transfer learning and to use pre-trained ones rather than the training of models from scratch. Many research institutions release models for large challenging datasets, and these pre-trained models can be used for extracting features from the new network. Because of the limited volume of data, this automated TB diagnosis problem embraces transfer learning as the solution and strives to explore, assess, and evaluate the impact of different CNN architectures.

There are two widely used methods to embrace pre-trained CNN's capabilities. The first technique does feature extraction where the extracted features are introduced into a new network to perform classification. This approach is widely used to maintain the valuable feature extractors trained during the initial stage. The second technique adds modifications to the pre-trained model that may include improvements to the architecture and parameter tuning. The basic information derived from the previous task is preserved, and new trainable parameters are introduced into the network. This is the technique opted for in the present study.

Various models like VGG-19 [18], DenseNet121 [19], ResNet50 [20], and InceptionV3 [21], which is already trained on ImageNet dataset, are utilized. After instantiating the model, pre-trained weights were loaded automatically, and images are resized to 128 * 128 with three channels without including top layers. The weights of the initial layers are made frozen to retrain only the higher layers to fine-tune the model. After several experiments, the parameters were established. Choices of parameters are possible and can be explored in the future for improving the results. After flattening the layer, a dense layer is added with arguments of 1024 neurons and the rectified linear unit (ReLU) [22] as an activation function. To avoid overfitting, dropout [23] of probability 0.4 is introduced. One more dense layer is added with 256 neurons with a dropout probability of 0.2. Subsequently, the dense layer is enabled by

Fig. 4 Flowchart


Start
Load Dataset
Prepare dataset, define constants and variables
Preprocess data by creating data generator and carry out data augmentation
Load pretrained model [VGG19, ResNet50, DenseNet121, InceptionV3]
Exclude top layer and freeze intended layers
Add new classifier layers
Define new model
Fit final model on training dataset and save the model
Make prediction on new images
Image detection
Yes
No
Categorize according to type (TB / Normal)
Does not find category
End

the softmax activation function to classify the image input as normal or infected. The learning rate of the model was set to 0.001. SGD [24] as the optimization method is used to compile CNN, and training was carried out using a categorical cross-entropy loss function of batch size 64. Early stopping was implemented to monitor the test loss at each epoch and to interrupt the training once the model improvement is stopped. The workflow is given, and a detailed process is included in Fig. 4 as a flowchart.

Procedure followed in flowchart:

1. Examine and prepare a dataset
2. Preprocess data by adding data augmentation using ImageDataGenerator class
3. Compose new model by loading the pre-trained model and add new classifier layers on top.
4. Train the model
5. Fit the final model on the training dataset and save it.
6. Generate predictions using the saved model: normal or tuberculosis.

2.2.3 Metrics

Specific metrics were recorded for CNN classification task which are as follows: (a) correctly TB cases recognized (TP, true positives), (b) incorrectly recognized TB cases (FN, false negatives), (c) correctly recognized no TB cases (TN, true negatives), and (d) incorrectly recognized no TB cases (FP, false positives). Based on these metrics, accuracy, sensitivity, and specificity of the model are calculated.

$$\text{Acurracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \tag{1}$$

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \tag{2}$$

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP}) \tag{3}$$

3 Results

The primary aim of using the transfer learning technique was to diagnose tuberculosis precisely in CXR images. For this, prepare and train all the models separately. In the Montgomery dataset, VGG-19 was trained, and it achieved ACC of 89% and AUC value was 88%. ResNet50 performs better than DenseNet121 and InceptionV3 in terms of accuracy and AUC. It achieved an AUC of 77% and an accuracy of 80%. As these measures are highly dependent on the number of samples representing each class, their subjective evaluation leads to incorrect conclusions. Because of this reason, the criterion for selecting the best model has to be the combination of sensitivity and specificity. Table 2 shows the achieved results for each CNN in terms of accuracy, sensitivity, and specificity. The best results were attained by the VGG-19 network.

In the Shenzhen dataset also, VGG-19 performed better than all other networks by achieving the best accuracy of 95% and AUC of 95% which is presented in Table

Table 2 Results of CNN with transfer learning-Montgomery dataset

Network used	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
VGG-19	89	98	77	88
ResNet50	80	91	65	78
DenseNet121	80	98	55	77
InceptionV3	70	83	51	67

Table 3 Results of CNN with transfer learning-Shenzhen dataset

Network used	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
VGG-19	95	96	93	95
ResNet50	85	79	91	85
DenseNet121	75	91	60	75
InceptionV3	73	93	43	74

3. The results attained in the Shenzhen dataset outperform the results achieved by the Montgomery dataset, probably due to the limited sample size and class imbalance of the samples. The Montgomery dataset has 60% samples negative with 40% positive, and in Shenzhen dataset, the sample class balance is almost 50% which adds as a favorable aspect.

VGG-19 outperformed in the combined dataset with the best results, followed by ResNet50 shown in Table 4. The accuracy and AUC achieved by VGG-19 are almost similar to Shenzhen dataset but vary in the ResNet50 model.

The confusion matrix of VGG-19 shown in Table 5 depicts true positives, true negatives, false positives, and false negatives of the model which helps to compare the best models further. The best outcomes are those with the lowest FN. A real-life perception of a false negative case will lead to the absurd conclusion that the patient is normal, which gives opportunities for public transmission of bacteria. Table 5 shows the low value of FN, especially for VGG-19, which is a good result. In terms of specificity also, VGG19 outperforms ResNet50 and other models and thus proves to be the most powerful model for the particular classification task in all the three datasets.

To visualize the performance of the model, the receiver operation characteristic (ROC) curve is plotted, as it is the most significant assessment metrics for testing

Table 4 Results of CNN with transfer learning-combined dataset

Network used	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
VGG-19	95	98	91	94
ResNet50	87	95	79	87
DenseNet121	74	97	50	74
InceptionV3	76	80	71	75

Table 5 Confusion matrix of VGG-19 and ResNet50

CNN	Dataset	TP	FP	FN	TN
VGG-19	Montgomery	79	13	1	45
	Shenzhen	316	21	10	315
	Combined	400	34	6	360
ResNet50	Montgomery	73	20	7	38
	Shenzhen	259	30	67	306
	Combined	386	82	20	312

the efficiency of any classification model. The ROC curve plotting true positive rate (sensitivity) against false positive rate (1 − specificity) of the VGG-19 model is presented in Fig. 5 for all the three datasets. It is a proven fact that the higher the AUC, the better the model is to differentiate between patients with TB and normal. Classifiers give curves closer to the top-left corner, indicating that a better accuracy and AUC were achieved.

4 Discussion

The findings of other papers that use the same dataset as ours were compared [25]. Proposed a CNN network to reduce computational as well as memory requirement and achieved 79%, 84.4%, 86.2% in terms of accuracy without pretraining the model [26] obtained 90.3% and 0.96 in accuracy and AUC with transfer learning after training a set of CXRs around 10,848 [27] proposed a model with 19 layers and showed an accuracy of 94.73% with validation accuracy of 82.09% using Adam optimizer. Though in other papers full data is not provided for assessment, the proposed model performed better in terms of accuracy even after training the model on a smaller size database (Table 6).

To achieve more validity to the findings in a clinical-based environment, it is ideal to train the system with more images. This work has faced the limitation of sufficient publicly accessible datasets, and the results obtained could be improved by using larger datasets.

5 Conclusion

This study aims to propose a deep learning approach with the transfer learning technique for classifying and identifying TB from CXR images. To extract features from X-rays, the pre-trained architectures VGG-19, RestNet50, DenseNet121, and InceptionV3 are used and trained on the ImageNet dataset. The approach of deep learning can be considered for bidirectional system validation in a clinical setting

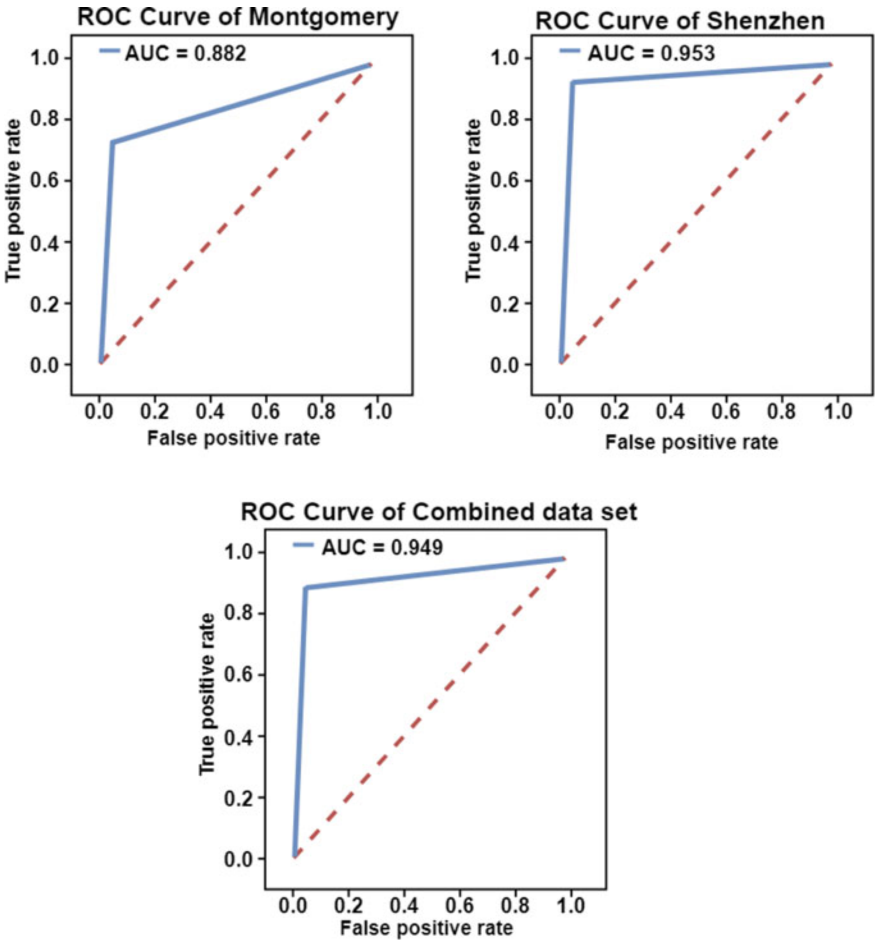


Fig. 5 ROC curves of VGG-19 on three different datasets. The AUC score is **a** 0.88 for dataset Montgomery, **b** 0.95 for dataset Shenzhen, **c** 0.94 for combined dataset

Table 6 Comparative results

Model	AUC (%)	Test accuracy (%)
Pasa [25]	92	86
Hwang [26]	96	90
Rahul Hooda [27]	–	94
Proposed approach	94	95

to detect TB by which both human and system error can be drastically minimized. This could be helpful for early diagnosis of TB, and to achieve this, the system must be trained with larger datasets. Thus it is recommended that high burden developing countries like India use this situation to share their chest X-ray dataset to the public, which may help to facilitate the development of much better tools.

References

1. Tuberculosis (2020) World Health Organization. https://www.who.int/news-room/fact-sheets/detail/tuberculosis
2. Global TB Report 2017. https://www.who.int/tb/features_archive/global_tb_report_2017/en/
3. Global Tuberculosis Report (2018) World Health Organization. https://apps.who.int/iris/btream/hadle/10665/274453/9789241565646eng.pdf
4. Shah S (2009) Intensified tuberculosis case finding among HIV-infected persons from a voluntary counseling and testing center in Addis Ababa, Ethiopia. J Acquired Immune Def Syndromes 537–545
5. Bakari M (2008) Basis for treatment of tuberculosis among HIV-infected patients in Tanzania: the role of chest X-ray and sputum culture.BMC 8
6. Harries AD, Maher D, Graham S (2004) TB/HIV: a clinical manual. WHO
7. Ndugga K, Klatser PR (2005) Sex-specific performance of routine TB DiagnosticTests. Int J Tuberc Lung Disease 294–300
8. WHO (2016) Chest radiography in tuberculosis detection. WHO
9. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press
10. Krizhevsky A, Hinton GE, Sutskever I (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems
11. Rouse M (2020) Techtarget. https://searchenterpriseai.techtarget.com
12. Bindhu V (2019) Biomedical image analysis using semantic segmentation. J Innov Image Process 91–101
13. Manoharan S (2019) Smart image processing algorithm for text recognition, information extraction and vocalization for the visually challenged. J Innov Image Process 31–38
14. Huang Z, Pan Z (2017) Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data. MDPI 9
15. Jaeger J (2014) Automatic tuberculosis screening using chest radiographs. IEEE Trans Med Imaging 33(2):233–245
16. Jaeger J (2013) Automatic screening for tuberculosis in chest radio graphs: a survey. Quant Imaging Med Surg
17. Weinberger J (2013) Principles of pulmonary medicine. Elsevier Health Sciences
18. Bansal S (2018) Kaggle. https://www.kaggle.com/shivamb/cnn-architectures-vgg-resnet-inception-tl
19. Keras. https://keras.io/api/applications/densenet
20. Dwivedi P (2019) Towards data science. https://towardsdatascience.com/understanding-and-coding-a-resnet-in-keras-446d7ff84d33
21. Keras. https://keras.io/api/applications/inceptionv3
22. Brownlee J (2020) A gentle introduction to the rectified linear unit. https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks
23. Brownlee J (2019) A gentle introduction to dropout for regularizing deep neural networks. https://machinelearningmastery.com/dropout-for-regularizing-deep-neural-networks
24. Keras. https://keras.io/api/optimizers/sgd
25. Pasa F, Golkov V, Pfeiffer F (2019) Efficient deep network architectures for fast chest X-Ray tuberculosis screening and visualization. Nature
26. Hwang S, Kim HE, Jeong J (2016) A novel approach for tuberculosis screening based on deep convolutional neural networks. In: SPIE medical imaging
27. Hooda R, Sofat S, Kaur S (2017) Deep-learning: a potential method for tuberculosis detection using chest radiography. IEEE

E-commerce Logistic Route Optimization Deciphered Through Meta-Heuristic Algorithms by Solving TSP

M. Soumya Krishnan and E. R. Vimina

Abstract E-commerce business is now becoming more popular and has become a critical need of the day for every customer. When considering the revenue and expense factors, each e-commerce business spent their major outflow in its logistics activities. Logistics costs are the focused area to be optimized so that the overall business can be dealt with in a controlled and safe manner. To with the logistics part, let us first consider the problem of vehicle routing. The Vehicle Routing Problem (VRP) is an intricate situation where multiple constraints can be taken into consideration. Thus, before dealing with such a complex situation, originally to move on with the study, the simple Traveling Salesman Problem is considered here in the paper. To solve such a hard NP problem and to get an optimal solution, many alternate methods have been tried out by many. In this paper, some of the meta-heuristic algorithms are considered and each of their performance is compared depending upon the length of the route, number of cities, time of execution of each algorithm, and their error rates. These algorithms are implemented and verified using Python 3.8.

Keywords Traveling salesman problem (TSP) · Logistic optimization · Meta-heuristics · Swarm intelligence · E-commerce · Artificial fish swarm algorithm (AFSA)

1 Introduction

Logistics costs are classified into many types. In this study, the outbound logistics cost is taken into consideration by taking into view the transportation activities. The workflow of the study can be depicted as shown in Fig. 1. The flow graph depicts the correct flow of work to follow based on the research point of view. The main concentration of the original research study is to optimize the logistic cost for an

---

M. Soumya Krishnan (✉) · E. R. Vimina
Amrita School of Arts and Sciences, Amrita Vishwa Vidyapeetham, Kochi, India
e-mail: soumyamahesh15@gmail.com

E. R. Vimina
e-mail: vimina.er@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_32

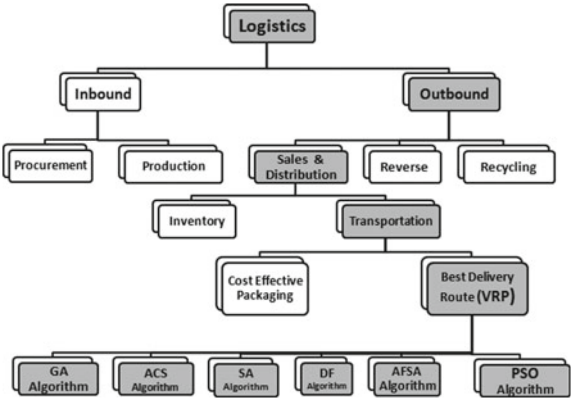


Fig. 1 Workflow of the study

e-commerce company whose major concentration is on outbound logistics. Even though other areas are involved in this, our study mainly focuses on the sales and distribution part. This is again managed by the transportation unit. When transportation is considered as the focus area of study, it was noticed from various literature reviews and other sources of study that VRP plays an imperative role in scheming logistics. Thus, VRP is taken as the main area of study based on which one can try to optimize the cost of logistics.

Some of the common objectives of a VRP can be projected out as follows minimize the total transportation cost based on the overall distance traveled as well as the costs allied with the drivers and the vehicles used, minimize the count of the vehicles required to serve all clients, calculate the travel time and capacity (vehicle load), and restrain penalties for poor quality service.

Also, VRP is a complex system where various variants as well as specializations are to be considered during the study. Some of the commonly considered VRP variants can be portrayed as VRP with pickup and delivery, with LIFO, with time windows, capacitated VRP, with multiple trips as well as open VRP. Along with these, multiple constraints need to be considered in solving each variant. Some of the possible constraints are shown in Fig. 2. Solving VRP by considered these multiple constraints is a complex process. But, one of the proven methodologies is by using various meta-heuristic algorithms. Thus, to start with the study, instead of the complex VRP, simple TSP is taken into consideration, using which, optimization of transportation cost can be done. When there's only one constraint such as a single-vehicle traveling through multiple cities is considered for the study, the VRP gets reduced to the TSP. Based on this concept; the study is performed using various meta-heuristics algorithms to get the most optimized result.

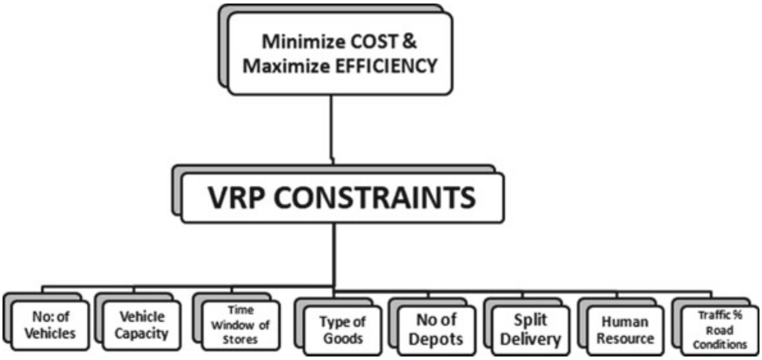


Fig. 2 Factors affecting VRP

1.1 Traveling Salesman Problem

Traveling Salesman Problem (TSP) is intended to explore the best possible roadmap for a traveler or a salesman where he tries to visit a set of cities based on the condition that every city should be touched exactly once, except the city where the travel was initiated and that must be the last city to visit. Such methodology is implemented in areas like logistics and transportation [1]. TSP is always well thought-out as an NP-hard problem in combinatorial optimization. Solving such problems using conservative methods is complex. This process of computations is liable to consume enormous time. Thus, try to compromise for approximated results which comparatively consumes less time, ending up with not an optimal solution but near to it.

Mathematically, TSP is described as follows:

$$\min T = \sum_{\mathrm{i}=1}^{n-1} d(\mathbf{x}_{\mathrm{i}}, \mathbf{x}_{\mathrm{i}+1}) \tag{1}$$

Here, X_i denotes the ith town or city in which i can be assigned values from 1 to n. Also, T is the total of the distance covered in the entire trip represented as $d(X_i, X_{i+1})$, in which X_i and X_{i+1} is the length between the town i and the next immediate $i + 1$th town [2].

The basic steps to solve a TSP problem can be illustrated as follows:

1. Generate a list of city coordinates
2. Initialize a function object using a coordinate list
3. Create a list of distances between pairs of cities
4. Initialize a function object using distance list
5. Define an object for the optimization problem.
6. Select and execute any chosen randomized optimization algorithm
7. Get the output as optimal route and minimal distance covered.

The steps mentioned above can be modified using different meta-heuristic algorithms.

TSP can be solved by various meta-heuristics approaches in a much efficient and faster method than any other proved customary methods by sacrificing aspects like optimality, speed, accuracy as well as precision [3]. Heuristics approaches have classifications like specific heuristics as well as meta-heuristics [1, 4]. Those heuristics which are used to solve specific problems are known as specific heuristics, whereas meta-heuristics are a certain class of algorithms that are used commonly to resolve almost any type of problems related to optimization [5]. One such popular meta-heuristic algorithm is swarm intelligence algorithms (SI), where a huge number of individual agents interact together and are used to portray the collective behavior of the system [6]. Some of the popular and widely used SI algorithms include the Genetic Algorithm, Artificial Bee Colony Algorithm, Ant Colony Optimization Algorithm, Differential Evolution Algorithm, Particle Swarm Optimization Algorithm, Cuckoo Search Algorithm, and so on.

The paper starts with an epigrammatic explanation on six SI-based algorithms as mentioned. Following that, experimentation is tried out to measure the performance of the prescribed algorithms based on the efficiency check and the time of execution. The results are conferred comprehensively along with the statistical analysis in the section followed. Based on the outcome, one best-performing algorithm is selected against the other five algorithms opted. The conclusion segment is projected at the end of this paper.

2 Related Work

In every e-commerce activity, it is found that the organization’s main expense is in its logistics. Thus, considering such a factor into consideration, in this paper, the problem is primarily considered as a TSP issue and the study is carried out with minimum constraints (as mentioned in the introduction part). In this paper, the efficiency of various swarm intelligence algorithms is compared on the basis of factors like the length of the trip and the total time taken for the same. Here, the datasets are generated randomly in which details like city coordinates are given in a pre-specified range from 10 to 200 or more. The datasets are retrieved from TSPLIB [7].

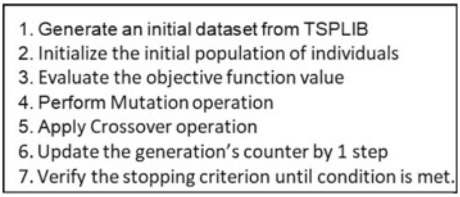
TSPLIB is a library of TSP sample instances where a lot of such samples are found in which varying sizes of dataset related to city counts and corresponding locations are found.

Here, similar problems from various sources are of different types. Each dataset in the library is categorized based on the number of cities. The attributes defined in each one are the type of data, dimension, tour section, and their corresponding (x, y) coordinates indicating the city positions. For our study purpose, the dataset opted are having a dimension of 10, 51, 100, and 200 cities.

In this experiment, the performance of optimization techniques selected is assessed on some factors like the shortest length of the route explored and the execution time of the algorithm. The experiment was implemented using the Python 3.8 version on a Sony VAIO Laptop, Core i5 processor, and operating system of version Windows 10. The algorithm was set to execute an average of around 30 iterations. The result was evaluated based on the mean value of the time taken (in a sec) to complete the execution process of each algorithm. Also, the efficiency is measured in terms of the minimal route length which each algorithm gave as the result based on the criteria of each algorithm set while coding using Python 3.8. If it is found that the mean value is less than 1.000E−10, then the result is stated as 0.000E+00. In this experiment, only the fundamental versions of swarm intelligence techniques are well thought-out and no specific modifications are introduced. Algorithm codes are tailored from multiple sources and updated to adhere to our experimental setup. In the forthcoming section, each SI algorithms taken for the study purpose is elaborated.

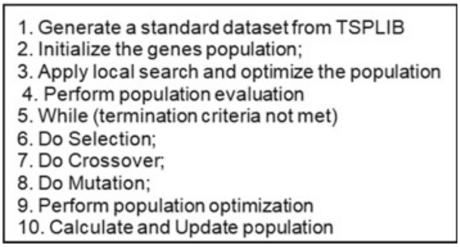
2.1 Differential Evolution (DE)

DE is a new type of progression algorithm which is alike GA. DE is a type of heuristic algorithm proposed by Kenneth Price and Dainer Storn in 1995 [2]. DE is a tool for optimizing a given population. It is used to obtain optimal solutions from a given set of constraints in various numerical problems [8]. DE is considered similar to the Genetic Algorithm as it has many similar basic operations. The mutation is one such to say. Similarly, crossover and selection are also operations of both. The fundamental difference between these two algorithms is in producing many improved solutions. When DE works on mutation, GA shows its expertise in crossover operation [6, 9]. DE works by introducing a mutant vector. During the mutation operation, it computes the difference in weights between two vectors which are selected randomly and then adding it to a third vector [2]. The procedure of the DE algorithm starts with *Initialization* followed by *Mutation*, then again the process of *Recombination,* and finally deals with *Selection.* The DE pseudocode for TSP [9] is explained using Fig. 3.



```
1. Generate an initial dataset from TSPLIB
2. Initialize the initial population of individuals
3. Evaluate the objective function value
4. Perform Mutation operation
5. Apply Crossover operation
6. Update the generation's counter by 1 step
7. Verify the stopping criterion until condition is met.
```

Fig. 3 Pseudocode of DE for TSP implementation



```
1. Generate a standard dataset from TSPLIB
2. Initialize the genes population;
3. Apply local search and optimize the population
 4. Perform population evaluation
5. While (termination criteria not met)
6. Do Selection;
7. Do Crossover;
8. Do Mutation;
9. Perform population optimization
10. Calculate and Update population
```

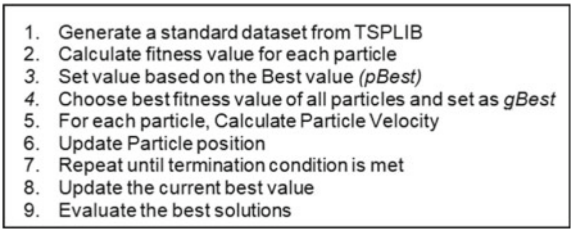
Fig. 4 Pseudocode of GA for TSP implementation

2.2 Genetic Algorithm (GA)

GA is a time-honored search-based heuristic optimization algorithm that operates on both the genetic strategy and the natural selection basis. The whole idea of GA was put forward by John Holland in the year 1975. The algorithm tries to mimic biological progression or evolution [6, 10]. GA is a specific type of algorithm that use techniques which are motivated by certain biological operations like inheritance. Mutation is another step in the process along with a selection followed by crossover or recombination. The working of the algorithm begins with finding problem solutions for the population of candidates. These candidate solutions are then evaluated on the basis of their problem-solving capability [11]. It is found that only the fittest survives. This then combines with each other, and finally, various possible solutions of next-generation are created. The algorithm is repeated and finds that this population of patterns “evolves” through operations like reproduction, mutation, and natural selection. The algorithm terminates when maximum numbers of generations are reached, or when it approaches a sensible fitness level for the considered population. The pseudocode of the Genetic Algorithm for TSP [10] is as shown in Fig. 4.

2.3 Particle Swarm Optimization Algorithm (PSO)

PSO again falls under the category of optimization algorithm enthused by swarm intelligence. In 1995, the algorithm was proposed by Kennedy and Eberhart. The algorithm is based on a simple mathematical model that tries to illustrate the communal behavior of birds, fishes, etc. The model relies mostly on the basic principles of self-organization and utilizes a simplified model based on social behavior to solve the optimization problems [11, 12]. PSO algorithm fundamentally works by taking into consideration a population or a group known as “*swarm of candidate solutions*” also identified as *particles*. This particle movement is restricted within the particular space of search. It works based on certain formulae to calculate the velocity of particle notated as (2) and also the position of the particle which is denoted by (3) [13]. Here, the best position of an individual particle is indicated as *pBest* and on the whole, the best position of the particle is recognized as *gBest*. These are then tracked



```
1. Generate a standard dataset from TSPLIB
2. Calculate fitness value for each particle
3. Set value based on the Best value (pBest)
4. Choose best fitness value of all particles and set as gBest
5. For each particle, Calculate Particle Velocity
6. Update Particle position
7. Repeat until termination condition is met
8. Update the current best value
9. Evaluate the best solutions
```

Fig. 5 Pseudocode of PSO for TSP implementation

by analyzing the activities of the particles in the swarm. Based on the improved positions, the movement of the swarm is guided forward [14]. The process continues until a satisfactory solution is discovered. The formula for the PSO algorithm is described as shown:

$$\boldsymbol{v}_{k+1}^{i} = \boldsymbol{v}_{k}^{i} + \boldsymbol{c}_1\boldsymbol{r}_1\left(\boldsymbol{p}_{k}^{i} - \boldsymbol{x}_{k}^{i}\right) + \boldsymbol{c}_2\boldsymbol{r}_2\left(\boldsymbol{p}_{k}^{q} - \boldsymbol{x}_{k}^{i}\right) \tag{2}$$

Gives the velocity of individual particles whose position is given by

$$\boldsymbol{x}_{k+1}^{i} = \boldsymbol{x}_{k}^{i} + \boldsymbol{v}_{k+1}^{i}. \tag{3}$$

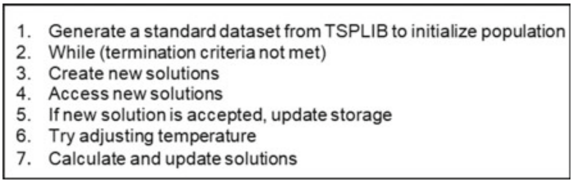
where

$\boldsymbol{x}_{k}^{\prime}$	Particle position
$\boldsymbol{v}_{k}^{i}$	Particle velocity
$\boldsymbol{p}_{k}^{i}$	Best "remembered" individual particle position (***pBest***)
$\boldsymbol{p}_{k}^{q}$	Best "remembered" swarm position (***gBest***)
c_1, c_2	Cognitive and social parameters
r_1, r_2	Random numbers between 0 and 1.

The PSO algorithm for TSP [12] is shown in Fig. 5.

2.4 Simulated Annealing (SA)

SA is found out as another effectual optimization algorithm that focuses on simulating the annealing of metals [15, 16]. Here, a solid metal's temperature is made to rise until it fuses in a heat bath. The metal is then slowly cooled until the particles are rearranged in the solid ground state. Thus, the metal's physical properties change with its internal structure. This will happen if the temperature is kept high enough, and it slowly decreases [17, 16]. The entire procedure is known as the simulation of the process of annealing. The method begins by having the heating cycle simulated with a temperature variable. A high initial value is attained to decrease the value gradually as the algorithm runs. Then, the algorithm is allowed to jump out of any local optimums during its execution. The opportunity to consider worse options will diminish as the temperature drops [17, 18]. Therefore, the algorithm is allowed to



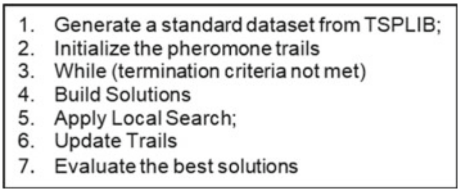
```
Generate a standard dataset from TSPLIB to initialize population
While (termination criteria not met)
Create new solutions
Access new solutions
If new solution is accepted, update storage
Try adjusting temperature
Calculate and update solutions
```

Fig. 6 Pseudocode of SA for TSP implementation

concentrate gradually on a region in the search space from which a near optimum solution is sought. Figure 6 depicts the pseudocode of SA for TSP [18].

2.5 *Ant Colony Optimization Algorithm (ACO)*

ACO is another SI-based algorithm that comes under the meta-heuristic category which was put forward by Marco Dorigo in the year 1992 [10, 19]. ACO works on the principle of indirect communication using pheromones, which is released by ants. Pheromones are the chemical substances that attract additional ants when looking for food. The magnetism of a given direction depends on the magnitude of the pheromones the ant detects. The excretions of pheromone follow some rules. It is not always the same strength that they display. Each quantity of excreted pheromones depends on the nature of the path they are navigating. Evaporation is the main pheromone mannerism, and the entire process is time-dependent. When a track is not used, these pheromones quickly evaporate and they start using some other route [20]. A colony or group of ants moves through diverse states of the problem which are influenced by two decision rules, namely *trails* and *attractiveness*. Also, they follow the other two mechanisms called *trail evaporation* and *daemon actions*. As mentioned, the main aim of the algorithm is to find the shortest path that is optimal based on the actions of ants searching for a path between the colony and the food source point. Thus, every single ant slowly builds a solution to the problem [11, 21]. Figure 7 depicts the pseudocode of ACO for TSP [20].



```
Generate a standard dataset from TSPLIB;
Initialize the pheromone trails
While (termination criteria not met)
Build Solutions
Apply Local Search;
Update Trails
Evaluate the best solutions
```

Fig. 7 Pseudocode of ACO for TSP implementation

2.6 Artificial Fish Swarm Algorithm (AFSA)

Another approach that is widely accepted in swarm intelligence is AFSA, which works based on stochastic search as well as the population under consideration [22, 23]. The school of fishes shows intellectual societal manners like random, searching, swarming, chasing, and leaping type of behaviors. The AFSO was first proposed in 2002. Since the system works on population, it is initialized first into a series of possibly randomly generated solutions and then iteratively searches for the optimum one [22]. The atmosphere in which the artificial fish lives is known as the solution space. Each fish's behavior depends on its current state and environmental situation. The environment is influenced by its own activities as well as the other companions' activities [22]. The three basic behaviors of AF are *prey* followed by the *swarm* and *then follow* [24]. Behaviors of each fish vary with the situation and are described as—*foraging, huddling, random, bulletin board,* and so on [25, 26].

The three basic behaviors and working of Artificial Fish can be illustrated [24] as:

(a) ***Prey_AF***: The fish perceives the concentration of food in water to assess the movement through vision or sensation and then selects the pattern [22, 26]
(b) ***Swarm_AF***: The fish must usually gather in groups during the moving cycle, which is a kind of living habits to guarantee the colony's survival and to escape dangers [26].
(c) ***Follow_AF***: The community partners will track and easily enter the food in the moving phase of the fish swarm when a single fish or multiple fish find food [24, 26].

The working of AFSA can be elaborated by considering the total number of AF as *N*, and its individual state as *X* whose range can be defined as $x_1, x_2, \ldots, x_n$, in which x_i gets the value from $1, 2, 3, \ldots, n$. These are the variables considered to be optimized. Next, assume that the best ever moving step of AF as *Step*. Also, consider the perceived_distance of the AF as *Visual*, δ to be the congestion_factor and the distance of the AF denoted as (i, j) as $\mathbf{d_{ij}} = \left|\mathbf{x_i} - \mathbf{x_j}\right|$. The concentration of food for the AF is depicted in terms of $Y = f(x)$. Here, the objective function value is taken as Y, and the number of tries is assumed as *Try_Num* [25, 26].

Similarly, the behavior of the fish differs from the situation. They can be described as follows:

2.6.1 Foraging _Behavior

This is a fundamental behavior of each AF, which is mentioned as the activity of moving toward the food. Here, the AF gets attracted to the food in water by perceiving the concentration of the food with the help of its vision [25, 27]. The current state of every AF is measured as x_i. Let x_j be a randomly chosen perception state which is given by:

$$x_j = x_i + \text{Visual} \cdot \text{Rand}() \tag{4}$$

Here, Rand( ) is any random number in the range between 0 and 1.

Also, consider the condition when $Y_i < Y_j$. Then, the step forward will be performed by AF in this direction notated as

$$x_i^{t+1} = x_i^t + \frac{x_j - x_i^t}{\left\| x_j - x_i^t \right\|} \cdot \text{Step} \cdot \text{Rand}(). \tag{5}$$

Else, repeat the try *Try_Num* times by randomly selecting status x_j to ensure whether the forward condition is satisfied. Still, if a forward condition is found dissatisfied, randomly select step by using (6).

$$x_i^{t+1} = x_i^t + \text{Visual} \cdot \text{Rand}(). \tag{6}$$

2.6.2 Huddling_Behavior

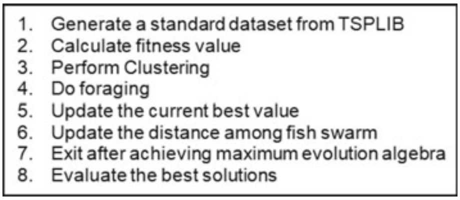
In this case, the fish will naturally seek to form a cluster to ensure the safety of groups and to avoid swimming hazards. AFSA says every fish should migrate as far as possible to the center of the neighboring partners and not be overcrowded [27]. For this, consider the current state of AF to be x_i. Next, assume that the search for the number of partners as n_f and x_c as its central location. If it is found that $Y_c/n_f > \delta Y_i$, it can be concluded that the partner at the center gets more food and the situation as not crowed. Thus, the forwarding step made toward the center of the partner is given by:

$$x_i^{t+1} = x_i^t + \frac{x_c - x_i^t}{\left\| x_c - x_i^t \right\|} \cdot \text{Stcp} \cdot \text{Rand}(). \tag{7}$$

If not, try performing the *foraging_behavior* again.

2.6.3 Following_Behavior

In this process, the neighboring partners will try swimming quickly to reach the point of food by following one fish that finds food. Rear-end behavior is just a chase behavior with the highest fitness for AF nearby. This process is also known as advancing toward a near-optimal partner. During this period, assume the current state of AF to be x_i. If x_j is set as the current neighborhood, then the maximum value of the partner to be searched is given as y_j. And if, $y_j/n_f > \delta y_i$, the situation is assumed as the state of x_j has a higher food concentration and is not found crowded. Thus, the forwarding step should be made toward x_j. Hence, it is given in (8) as shown:



```
1. Generate a standard dataset from TSPLIB
2. Calculate fitness value
3. Perform Clustering
4. Do foraging
5. Update the current best value
6. Update the distance among fish swarm
7. Exit after achieving maximum evolution algebra
8. Evaluate the best solutions
```

Fig. 8 Pseudocode of AFSA for TSP implementation

$$x_i^{t+1} = x_i^t + \frac{x_j - x_i^t}{\left\| x_j - x_i^t \right\|} \cdot \text{Step} \cdot \text{Rand}(). \tag{8}$$

Otherwise, repeat the steps of foraging behavior.

Some other behaviors exhibited by the AF are illustrated as follows:

2.6.4 Random_Behavior

This is usually considered as a default behavior of foraging. This action is to randomly pick a location, and then switch to the direction chosen.

It is meant for recording the individual state of the optimum AF position and the concentration of food. In this procedure of optimization, the own state as well as the billboard status measured per activity is checked for each individual AF. If the own state is found to be better than the status of a bulletin board, translate the billboard's own status and record the history of its optimal state.

This is used to record the autonomy behavior of every AF. When used for TSP implementation, it can be utilized to carry out cluster as well as to show rear-end behavior. Later, the same is used to evaluate the value of actions by selecting the optimal behavior to detect the value to be executed. Foraging_behavior is again found as its default behavior.

The pseudocode of AFSA for TSP [25] is given in Fig. 8.

These are considered as the basic behavior of a normal Artificial Fish Swarm Algorithm. Improved versions of AFSA [22] are also depicted in many other papers. But for our purpose of study, only local AFSA is used as per the explanation done in this paper.

3 Proposed Work Along with Its Experimental Results

The various meta-heuristic algorithms mentioned and discussed in this paper are implemented with a few alterations in some parameters in order to acclimatize and to decipher the problem of a traveling salesman. The investigational situation is executed using Python 3.8 programs implemented using a Sony VAIO Laptop, Processor is Intel Core i5, and the operating system is Windows 10. The TSP data used here are downloaded from the TSPLIB [7], and four test cases of a different number of cities were considered for the test to measure the performance of each meta-heuristic algorithm used [28]. These algorithms were iterated until they showed a congregated return value. An algorithm is usually said to converge when it shows the same result which can be considered best in its last 20–25 iterations. Here, the algorithms were tested by iterating it until their termination criteria were satisfied.

Figure 9 projects the route length comparison, and Fig. 10 shows the execution time taken by the six meta-heuristic algorithms referred here by considering a sample value of 51 cities from the dataset ***eil51***. It is clear that out of the six algorithms used in a sample dataset of 51 cities, the AFSA algorithm gives the optimal result with the least route length of 570.51. And comparatively the worst performance is shown by the SA algorithm. The result also shows that the AFSA algorithm again beats other algorithms in the execution process by considering 30 iterations with an average value of 2.283.

Fig. 9 Comparing route length obtained from different algorithms using eil51 dataset

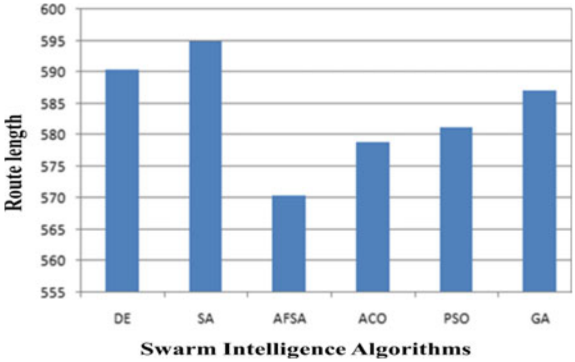
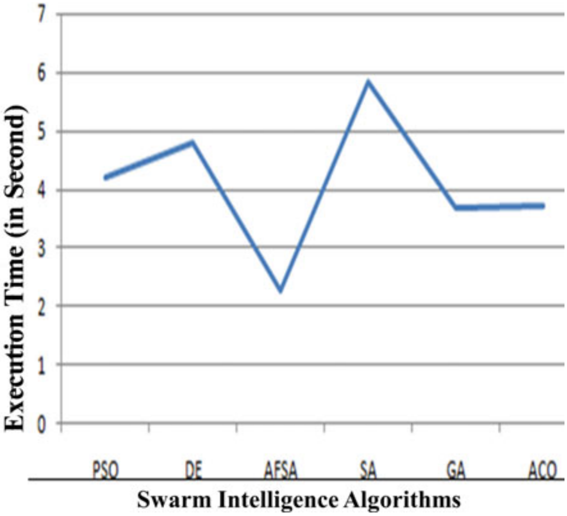


Fig. 10 Comparing the computation time of different algorithms using **eil51** dataset



During implementations, two main factors have to be considered to achieve better results. Locally searching within the problem space known as exploration and searching around the intended solution called extraction. Better results are achieved when one could balance these two factors. In the ACO algorithm, parameters such as α and β help to control the extraction of the algorithm which are empirically chosen as 1 and 2, respectively. In the SA algorithm, a high temperature is considered first by setting the exploration to large and as temperature decreases the extraction gradually becoming more. Therefore, first temperature T is considered to be high and then based on that, a cooling function is designed, which helps to lower the temperature slowly to obtain a high extraction. In the PSO algorithm, exploration and extraction are controlled by changing the weight of matrices accordingly. In the GA algorithm, crossover operation is used to increase exploration, and mutation operation helps for a better solution by implementing a local search on a solution by performing extraction [29, 30]. In the AFSA algorithm, the step size of the artificial fish and the number of the feeding process indicated as try-num helps to control the exploration and extraction. This adjustment is clearly mentioned in AFSA in Sect. 2 of this paper. Finally, with the DE algorithm, to control the exploration and extraction, it requires the six fundamental factors such as step size function indicated as *scaling factor*, *population number, initialization, crossover probability, evaluation function, and termination.*

Table 1 picturizes the performance or efficiency comparison of the different metaheuristic algorithms mentioned here based on the four datasets selected. They are averaged and set over 30 runs. The corresponding time of execution of each algorithm was also measured based on a particular dataset of 51 cities. The result was evaluated based on the mean value of the time taken (in a sec) to complete the execution process of each algorithm. And the error rate of each was also noted. If it is found that the mean value is less than 1.000E−10, then the result is stated as 0.000E+00. Likewise, the output of each algorithm was taken, and the best case considering the dataset ***eil51*** was considered for finalizing the output. The corresponding results are displayed here within Table 2.

From the output received, it is very evident that an AFSA algorithm shows a better result than others when considering certain factors as mentioned. Thus, the Artificial Fish Swarm Algorithm works better than other algorithms mentioned here in this paper when considering a sample TSP dataset of 51 cities. The termination criteria considered here are the number of iterations which are manually set to 30. Tables 1 and 2 show the results so obtained by the algorithms.

Table 1 Performance comparison of algorithms based on route length

TSP Dataset	No. of Cities	AFSA	GA	ACO	PSO	DE	SA
burma10	10	271.628	300.75	319_95	325.69	342.12	321.28
eil51	50	570.515	537.18	578.8	581.24	590.47	594.85
rd100	100	801.746	847.57	809.51	837.52	874.56	869.63
KroA200	200	3007.48	3218.04	3092.96	3146.36	3345.28	3324.27

Table 2 Performance comparison of various algorithms based on its execution time, route length, and error rate for a sample dataset of 51 cities (eli51)

Meta-heuristic algorithms	Execution time	Route length	Error rate
AFSA	2.283	570.51	4.09E−01
GA	3.673	587. 18	1.78E−15
ACO	3.727	578.8	4.38E−01
PSO	4.214	581.24	8.21E−08
DE	4.805	590.47	2.00E+00
SA	5.821	594.85	1.37E−10

The output results represent the performance of each algorithm based on the selected dataset eli5, which is a dataset of 51 cities taken as an average among other datasets used for the study. The figures show the performance of each algorithm based on its route length and its execution time. Figure 11 projects the efficiency of the ACO algorithm based on its execution time against the number of iterations as well as the route length obtained against the set number of iterations.

Figure 12 emphasizes the efficiency of GA based on the route length and execution time against the number of runs of the algorithm. Figure 13 points out the performance

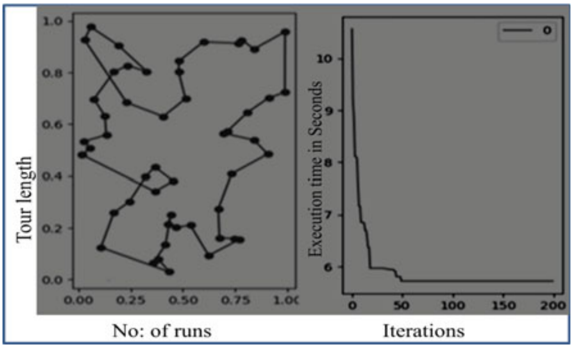


Fig. 11 Performance of **ACO** based on eli51

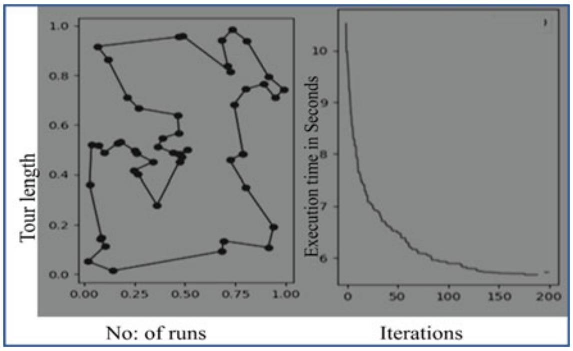


Fig. 12 Performance of **GA** based on eli51

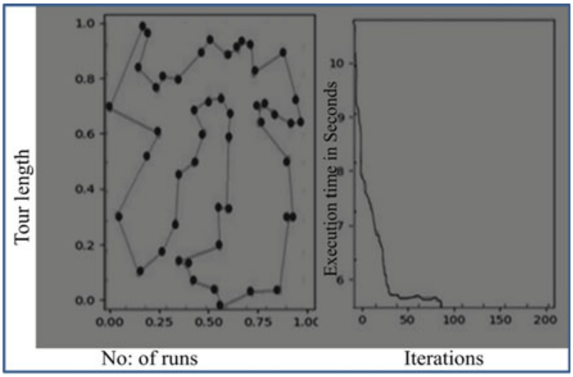


Fig. 13 Performance of **AFSA** based on eli51

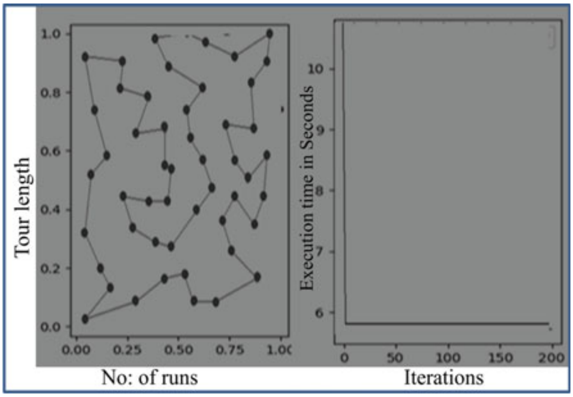


Fig. 14 Performance of **DE** based on eli51

of AFSA based on execution time and route length set against the iterations performed on the algorithm. Figure 14 represents the performance of DE algorithms based on the number of times the algorithm is executed against the time it takes to execute and produce the result along with the optimized route length. Similarly, Fig. 15 depicts the performance of PSO, and Fig. 16 showcases the efficiency of SA based on the number of iteration set to 30 against the route length and the execution time. Based on these findings, it can be concluded from the results that the AFSA algorithm is better than the other five algorithms used herein the analysis, both in terms of the length of the resulting routes as well as in terms of their execution time. AFSA is thus proved to have the advantages of fast execution speed as well as a strong ability to show the global convergence.

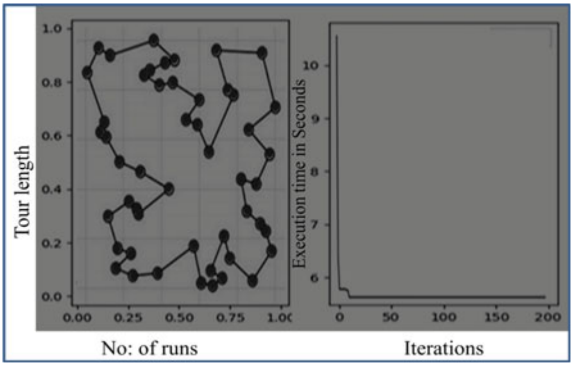


Fig. 15 Performance of **PSO** based on eli51

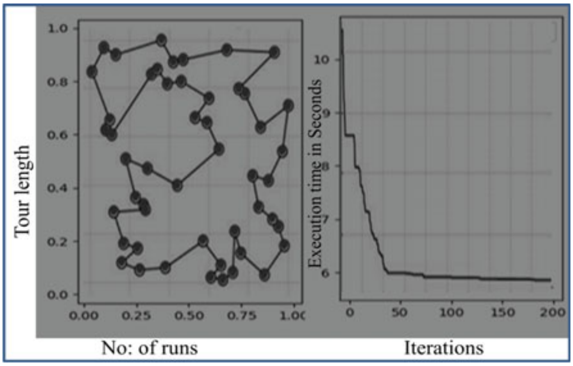


Fig. 16 Performance of **SA** based on eli51

4 Conclusion

The study was initiated to find an optimal solution to minimize the overall logistic cost incurred by an e-commerce firm. Thus, to move on with the work, it was decided to make minor improvements in the Vehicle Routing Problem (VRP) which can create a huge difference in the overall logistic cost. But, VRP problems deal with multiple constraints and factors. Initially, TSP was taken into consideration where it was dealt with only a single constraint "Number of Cities." But, finding an optimal route for a TSP is also critical to saving time and cost. In this work, six meta-heuristic swarm-based algorithms were considered and implemented to discover the finest minimum length of the route for the problem of traveling salesman. For this, four random datasets were selected from the TSPLIB consisting of cities from 10 to 200 in numbers. Each of the algorithms was individually tested to measure its performance based on the optimal route length along with its execution time. The set of algorithms considered for the study are the Genetic Algorithm, Simulated Annealing, Ant Colony Optimization, Artificial Fish Swarm Algorithm, Particle Swarm Optimization, and the Differential Evolution. These were implemented using Python 3.8 to measure the individual performances of these approaches. The results showed the superiority of AFSE with the ability to outperform others by covering the number of cities with minimum route length as well as with its minimum execution time of the algorithm to conclude with the result. The results identified AFSA as the best-performing algorithm to solve the TSP followed by ACO, then GA, PSO, SA, and finally DE. Every algorithm has got its advantages and shows its best performances which are selected based on the condition. Though AFSA demonstrates smarter behavior and produces more efficient performance than other swarm intelligence algorithms, it has got some shortcomings like showing a fall in local optimum points and advance in convergence in many situations while using multiple constraints. The conclusion is that the AFSA algorithm is one of the best swarm intelligence techniques, with key advantages like high convergence speed, versatility, resistance to errors, and high precision. Thus, researchers can use this along with its modified versions for their future research which involves multiple constraints such as in Vehicle Routing Problems in e-commerce to optimize the logistic cost.

References

1. Shima S, Mohammad S, Fardad F (2016) A comparison between swarm intelligence Algorithms for routing problems. Electr Comput Eng Int J (ECIJ) 5(1):17–33
2. Mei M, Xue H, Zhong M, Gu Y (2010) An improved differential evolution algorithm for TSP problem. In: International conference on intelligent computation technology and automation, pp 544–548
3. Vincent K, Matthew N, Spencer S (2014) Heuristics algorithms. (ChE 345 Spring)
4. Zharfi V, Mirzazadeh A (2013) A novel metaheuristic for travelling salesman problem. J Ind Eng 2013, Article ID-347825:5
5. El-ghazali T (2009) Meta-heuristic from design to implementation. Wiley Inc Publications, NY
6. Ab Wahab MN, Nefti-Meziani S, Atyabi A (2015) A comprehensive review of swarm optimization algorithms. PLoS ONE 10(5):e0122827
7. TSPLIB datasets http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/tsp/
8. Vanita GT (2013) Travelling salesman problem using differential evolutionary algorithm. IOSR J Eng (IOSRJEN) ISSN (e): 2250-3021, ISSN (p): 2278-8719, pp 63–67
9. Xiang W, Guoyi X (2011) Hybrid differential evolution algorithm for traveling salesman problem. Adv Control Eng Inf Sci, Procedia Eng 15:2716–2720
10. Sabry AH, Benhra J, El Hassani H (2015) A performance comparison of GA and ACO applied to TSP. Int J Comput Appl (0975–8887) 117(1):28–35
11. Hosam HAM, Ashraf YAM (2015) Performance comparison of simulated annealing, GA and ACO applied to TSP. Int J Intel Comput Res 6(4):647–654
12. Appiah MY, Xiong Q (2019) Route optimization in logistics distribution based on particle swarm optimization. Int J Comput Appl (0975–8887) 178(30)
13. Yudong Z, Shuihua W, Genlin J (2015) A comprehensive survey on particle swarm optimization algorithm and its applications. Math Prob Eng 2015, Article ID 931256:38
14. Thirachit S (2018) Enhancing particle swarm optimization using opposite gradient search for travelling salesman problem. Int J Comput Commun Eng 7(4):167–177
15. Ai-Hua Z, Li-Peng Z, Bin H, Song D, Yan S, Hongbin Q, Sen P (2019) Traveling-salesman-problem algorithm based on simulated annealing and gene-expression programming. MDPI Inf J

16. Sumathi M, Rahamathunnisa U, Anitha A, Druheen D, Nallakaruppan. MK (2019) Comparison of particle swarm optimization and simulated annealing applied to travelling salesman problem. Int J Innov Technol Explor Eng 8(6):1578–1583 ISSN: 2278-3075
17. Xiutang G, Zehui S (2009) An effective simulated annealing algorithm for solving the traveling salesman problem. J Comput Theoret Nanosci 6:1680–1686
18. Ai-Hua Z, Li-Peng Z, Bin H, Song D, Yan S, Hongbin Q, Sen P (2019) Traveling-salesman-problem algorithm based on simulated annealing and gene-expression programming. Information 10(7)
19. Sapna K, Ibraheem AQA (2015) Ant colony optimization: a tutorial review. In: Conference paper-national conference on advances in power and control, at faculty of engineering and technology, International University, Haryana
20. Ivan B Jr, Zuzana Č (2011) Solving the travelling salesman problem using the ant colony optimization. Manag Inf Syst 6(4):010–014
21. Hui Yu (2014) Optimized ant colony algorithm by local pheromone update. TELKOMNIKA Indo J Electr Eng 12(2):984–990
22. Mehdi N, Ali A, Ghodrat S, Mehdi S, Adel NT (2012) A review of artificial fish swarm optimization methods and applications. Int J Smart Sens Intell Syst 5(1): 107–148
23. Nitesh MS (2020) Solving random travelling salesman problem using firefly algorithm. Int J Innov Technol Explor Eng 9(4):1037–1041. ISSN: 2278-3075
24. Yun C (2010) Artificial fish school algorithm applied in a combinatorial optimization problem. I.J. Intell. Syst. Appl 1:37–43
25. Teng F, Liyi Z, Yang Li, Yulong Y, Fang W (2014) The artificial fish swarm algorithm to solve traveling salesman problem. Adv Intell Syst Comput 679–685
26. Nitesh MS, Sanjay PP (2020) Solving a combinatorial optimization problem using artificial fish swarm algorithm. Int J Eng Trends Technol 68(5):27–32. ISSN: 2231-5381
27. Nurezayana Z, Azlan MZ, Safian S (2015) Overview of Artificial Fish Swarm Algorithm And Its Applications In Industrial Problems. Appl Mech Mater Trans 815:253–257
28. Haider AA, Ibrahim FA (2015) Comparison of algorithms for solving traveling salesman problem. Int J Eng Adv Technol 4(6):76–79. ISSN: 2249–8958
29. Gamal Abd ENS, Abeer MM, El-Sayed MEl-H (2014) A comparative study of meta-heuristic algorithms for solving quadratic assignment problem. Int. J Adv Comput Sci Appl 5:1
30. Elham D, Arash M (2017) Meta-heuristic approaches for solving travelling salesman problem. Int J Adv Res Comput Sci 8(5). ISSN No. 0976-5697

A Comparative Study on the Performance of Deep Learning Algorithms for Detecting the Sentiments Expressed in Modern Slangs

Vivank Sharma, Shobhit Srivastava, B. Valarmathi, and N. Srinivasa Gupta

Abstract Sentiment analysis is a text investigation technique that distinguishes extremity inside the text, regardless of whether an entire document, sentence, etc. Understanding individuals' feelings are fundamental for organizations since customers can communicate their considerations and emotions more transparently than any other time in recent memory. In this paper, the proposed model is the sentimental analysis on Twitter slangs, i.e., tweets that contain words that are not orthodox English words but are derived through the evolution of time. To do so, the proposed model will find the root words of the slangs using a snowball stemmer, vectorizing the root words, and then passing it through a neural network for building the model. Also, the tweets would pass through six levels of pre-processing to extract essential features. The tweets are then classified to be positive, neutral, or negative. Sentiment analysis of slangs used in 1,600,000 tweets is proposed using long short-term memory (LSTM) network, logistic regression (LR), and convolution neural network (CNN) algorithms for classification. Among these algorithms, the LSTM network gives the highest accuracy of 78.99%.

Keywords Snowball stemmer · Classifiers · Sentiment analysis · LSTM · CNN · Logistic regression · Twitter · Word2Vector

---

V. Sharma · S. Srivastava
Department of Information Technology, Vellore Institute of Technology, Vellore, Tamil Nadu, India
e-mail: vivanksharma@ymail.com

S. Srivastava
e-mail: shobhit.sri0108@gmail.com

B. Valarmathi (✉)
Department of Software and Systems Engineering, School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India
e-mail: valargovindan@gmail.com

N. Srinivasa Gupta
Department of Manufacturing, School of Mechanical Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India
e-mail: guptamalai@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_33

1 Introduction

The Web has significantly changed how individuals express their perspectives and assessments. Sentiment analysis refers to the task of recognizing suppositions, positivity decisions, and other data identified with the emotions and frames of mind communicated in ordinary language writings. Some of the emotions are happy, sad, frustrated, angry, and so on.

Slang is the utilization of exceedingly casual words, shortened forms, and articulations that are dismissed when asked to be taken as a significant aspect of the traditional language. The flood of online communications, for example, electronic mail, text messages, and microblogs administration made use of Internet slang practically omnipresent. It has turned out to be critical to gauge feeling extremity of the opinion or sentiment arranged slang present on the Internet that shows up in the investigation or analysis. Opinion analysis attempts to choose the sort of opinion (affirmative or not) in a given content. Emotion or feeling-based classification has a couple of basic characteristics, including different tasks, highlights, strategies, and context. A new strategy is proposed to classify the sentiment expressed in the microblog slangs in this paper. At that point, the proposed model offers a technique to aid in deciding the extremity available in Slangs.

2 Related Works

Almost all existing sentiment analysis algorithms to date are designed in such a way, that it classifies any content into two binary class, i.e., either it is positive or else negative [1–3]. Few recently emerged algorithms were designed in such a way that it extends binary classification to multipoint rating system, i.e., rating inference or multi-category problem [4–6]. Recently few researchers have proposed models for analyzing the reviews of a product or service with respect to all its features [7, 8]. For slangs, a slang dictionary is used to calculate the TF-IDF score to decide the polarity [9]. They made an underlying stride at programmed identification and recognizable proof of slang from normal sentences utilizing deep learning techniques. They showed how phonetic highlights joined with deep learning algorithms offer interpretability. They found that the bidirectional LSTM with feature-based inputs and character-based convolutional embeddings utilizing multilayer perceptron yields the best performance in position recognizable proof, and the model with comparative components aside from with conditional random field has better execution in distinguishing regardless of whether a source sentence contains a slang term [10]. Proposed Twitter information to identify depression [11].

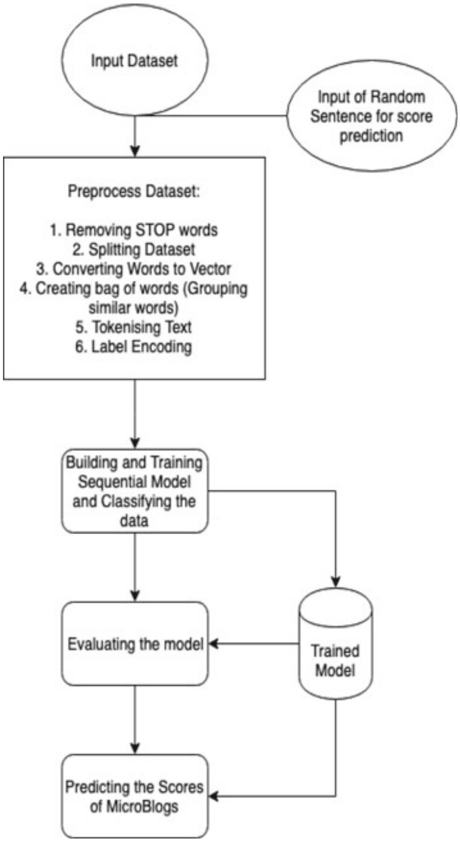
Better data pre-processing methods like changing over emojis to message structure, changing overstretched words to normal form, etc. were utilized to improve the identification accuracy. They extracted the highlights utilizing BOW, TF-IDF with n-grams, and Word2Vec procedures and used these highlights to the methods of classification. Logistic regression classifier (TF-IDF with n-grams) provided a 81% of the most extreme accuracy. Proposed a novel method called Representative Term-Document Matrix (RTDM) [12]. The given text document was transformed into a vector consisting of eight terms like bad, very bad, disgusting, never recommended, good, very good excellent, and recommended. A classification method is described using the Mahalanobis Distance (MD) [13]. The classifier name was Mahalanobis Distance Classifier (MDC). For 25,000 movie reviews, MDC achieved a 70.8% of accuracy. The hybrid classifier (MDC + MLP) performed with 98.8% of accuracy for 25,000 movie reviews. Deep learning technique is utilized to figure out sentiment analysis issues, for example, sentiment polarity [14]. Models utilizing term frequency-inverse document frequency (TF-IDF) and word embedding were applied to various datasets like Sentiment140, IMDB Movie Reviews, Cornell Movie Reviews, etc. CNN, DNN, and RNN were used in thisarticle.

3 Proposed Method

Current algorithms for sentimental analysis work great on formal English literature, but it fails when it comes to slangs. As slangs have no definite list and have no exact meaning, it is all based on contextual and scenario. So, the proposed model, a new algorithm, is developed which not only focuses on keywords but also takes into consideration every single word and tries to overcome slang by making custom datasets and using stemmers to find its original keyword and meaning. Hence, it makes our model different in the form and capability that it can also detect and analyze informal English, which is mostly used in microblogs. For example, for love, one can write many variants which are not in the English dictionary like "Luv," "Lub," etc. and the same one word can be even used to describe sentences also. "Shoulda" is replaced by should have. Similarly "lol" is replaced by "laugh out loud." So, to overcome this, the proposed model is taking more than 1 million Twitter data into account and training our model over these slangs and hence predicting it. Our model is achieving accuracy on the range of 70–80% on different datasets of different microblogs. The proposed model analyzes the microblogs by calculating a score by considering every word and slang and a sentiment score is calculated on a scale of 0-1 to assign sentiment lable. This can aid an organization to better comprehend the social sentiment of their service or their product by analyzing online conversations.

The first step is to preprocess the dataset. The second step is building and training the sequential model. The third step is to evaluate the model. The next step is to predict the score of Microblogs. The proposed model is summarized in the flowchart given in Fig. 1. The sentiment 140 dataset can be found in the link https://gofile.io/d/YS7U1r. It consists of 1,600,000 tweets that were obtained through the Twitter API. 1,280,000 tweets are used for the training and 320,000 tweets are used for testing. In the dataset, the class label "0" represents a negative tweet, whereas "4" represents a positive tweet. This can be used to detect sentiments [15]. The dataset label distribution which is depicted in Fig. 2 displays that both the positive and

Fig. 1 Proposed work flow chart


Input Dataset
Input of Random Sentence for score prediction
Preprocess Dataset:
1. Removing STOP words
2. Splitting Dataset
3. Converting Words to Vector
4. Creating bag of words (Grouping similar words)
5. Tokenising Text
6. Label Encoding
Building and Training Sequential Model and Classifying the data
Trained Model
Evaluating the model
Predicting the Scores of MicroBlogs

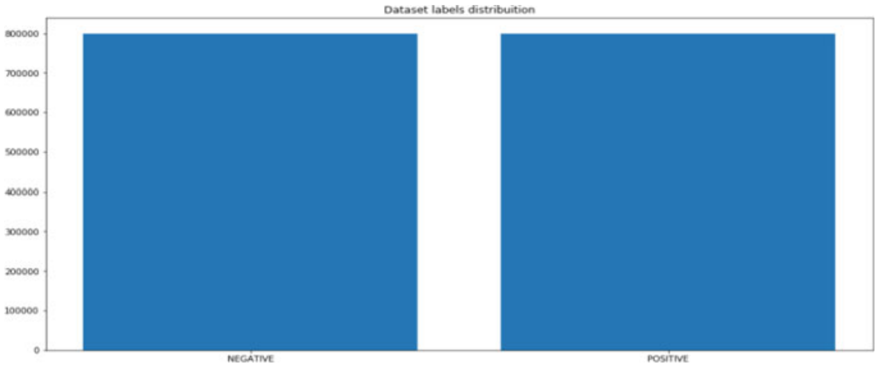


Fig. 2 Dataset distribution

negative labels are almost equal to the preprocessed data by removing the stop words. Content may contain stop words like "the," "is," "are." Stop words can be separated from the material to be prepared. There is no complete rundown of stop words in natural language processing inquire about; anyway, the NLTK module holds a group of stop words. All the stop words present in the dataset are removed. Then, the proposed model has used snowball stemmer for reducing inflected words in the dataset to their word stem, base, or root form. A stemmer is a calculation that works on the guideline of perceiving "stem" words implanted as such. These are useful for lexical purposes, for instance, in online lexicons, for heuristics in a record the board, or anyplace else that semantic apparatuses can help make a request.

Stemmers get the consideration of a center or stem word inside a more drawn out term. For instance, a stemming calculation may take a gander at a name like "planning," and accurately perceive that the root word or stem word is "plan." This can be a useful component of something that parses crude content for investigation, either for a site or some other venture. Snowball stemmer is famous for its compatibility with slangs. The next step is to convert all the text to vectors. These vectors are used by deep learning models over millions of words. The Word2Vec model helps in considering the context for further processing.

Word2Vec is a gathering of models which infers relations between a word and its relevant words. The two significant models inside Word2Vec are skip gram and continuous bag of words (CBOW).

The proposed model takes an inside word and a window of setting (neighbor) words and the model endeavor to foresee setting terms out to some window estimate for each middle name in skip gram model, Along these lines, our model will characterize a likelihood appropriation, for example, the likelihood of a word showing up in the setting given an inside word, and the next step will pick our vector portrayals to expand the possibility.

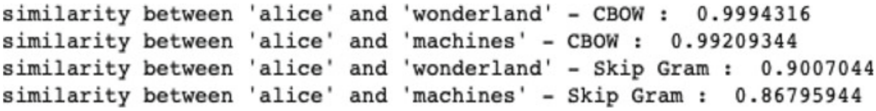
CBOW is just the inverse of the skip gram model. Try to forecast core-word by adding, vectors of neighboring words.

Skip gram as well as CBOW model output is shown in Fig. 3.

An example of a data pre-processing model is given in Fig. 4.

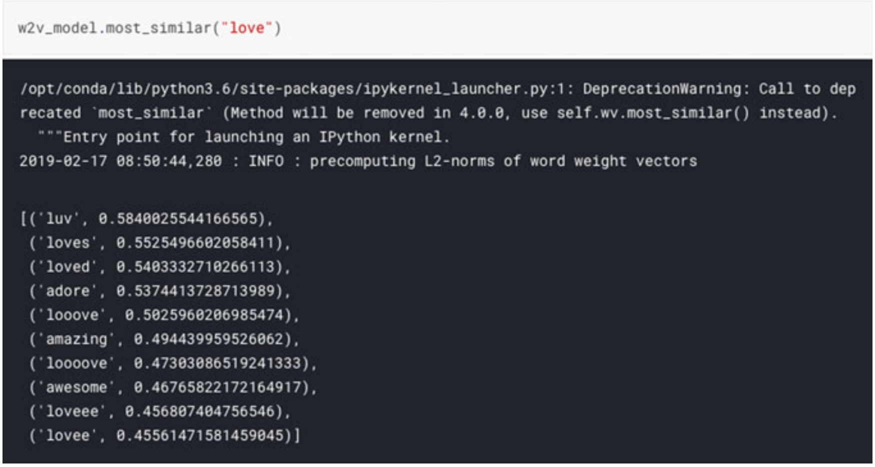
After this, tokenize the text, and then label encodes it for the further training process. Now, built the sequential model and it consists of four layers. Figure 5 shows the sequential model summary.

In the proposed model, the sigmoid activation function is used for the dense hidden layer. A sigmoid function exists between (0–1) and hence is the best model for predicting probability. The formula for the sigmoidal function is shown in Eq. (1).



```
similarity between 'alice' and 'wonderland' - CBOW :  0.9994316
similarity between 'alice' and 'machines' - CBOW :  0.99209344
similarity between 'alice' and 'wonderland' - Skip Gram :  0.9007044
similarity between 'alice' and 'machines' - Skip Gram :  0.86795944
```

Fig. 3 Example of data processed with skip gram as well as CBOW model

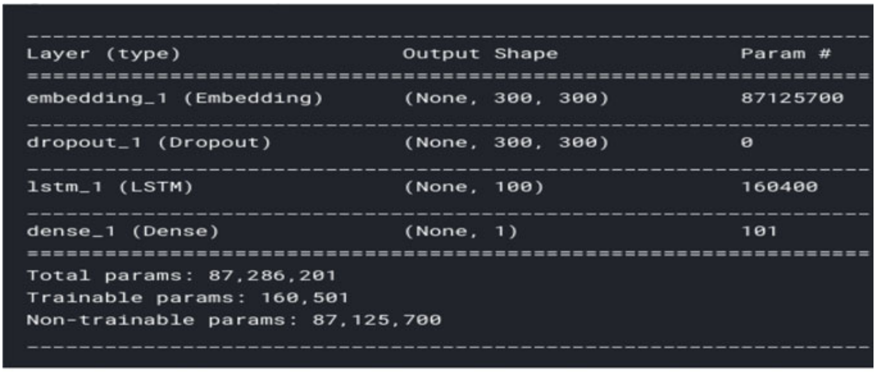


```
w2v_model.most_similar("love")
```

```
/opt/conda/lib/python3.6/site-packages/ipykernel_launcher.py:1: DeprecationWarning: Call to dep
recated `most_similar` (Method will be removed in 4.0.0, use self.wv.most_similar() instead).
  """Entry point for launching an IPython kernel.
2019-02-17 08:50:44,280 : INFO : precomputing L2-norms of word weight vectors


[('luv', 0.5840025544166565),
 ('loves', 0.5525496602058411),
 ('loved', 0.5403332710266113),
 ('adore', 0.5374413728713989),
 ('looove', 0.5025960206985474),
 ('amazing', 0.494439959526062),
 ('loooove', 0.47303086519241333),
 ('awesome', 0.46765822172164917),
 ('loveee', 0.456807404756546),
 ('lovee', 0.45561471581459045)]
```

Fig. 4 Data pre-processing grouping similar words



```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
embedding_1 (Embedding)      (None, 300, 300)          87125700
_________________________________________________________________
dropout_1 (Dropout)          (None, 300, 300)          0
_________________________________________________________________
lstm_1 (LSTM)                (None, 100)               160400
_________________________________________________________________
dense_1 (Dense)              (None, 1)                 101
=================================================================
Total params: 87,286,201
Trainable params: 160,501
Non-trainable params: 87,125,700
_________________________________________________________________
```

Fig. 5 Sequential model summary

$$\sigma(x) = 1/(1 + \exp(-x)) \tag{1}$$

Now, finally train the proposed model on the dataset, which consists of 1,600,000 tweets that were obtained through the Twitter API. In the dataset, the class label “0” represents negative tweets, whereas “4” represents positive tweets. This can be used to detect sentiments [15]. So, that it can further be used for predicting scores of the content to be analyzed.

4 Result and Discussion

On evaluating the test dataset, the proposed model gets an accuracy of 78.99% (LSTM Network), which is excellent compared to earlier achieved accuracy of 56.58–76.69% considering slangs.

Dang et al. [14] used sentiment140 dataset. Accuracy of DNN, CNN, and RNN algorithms with TF-IDF was 76.50%, 76.69, and 56.58. Among these algorithms, CNN was performing well. In the proposed model, the LSTM network, logistic regression, and CNN algorithms are used for processing. Among these algorithms, LSTM network was performing well. LSTM network gets an accuracy of 78.99%. Figure 6 shows the training, as well as the validation accuracy graphs.

Figure 7 shows the training and validation loss graph.

The confusion matrix is shown in Fig. 8. Hence from the confusion matrix, true positive, false positive, true negative, and false negative values are shown as true positive is 0.77, false positive is 0.19, true negative is 0.81, and false negative is 0.23.

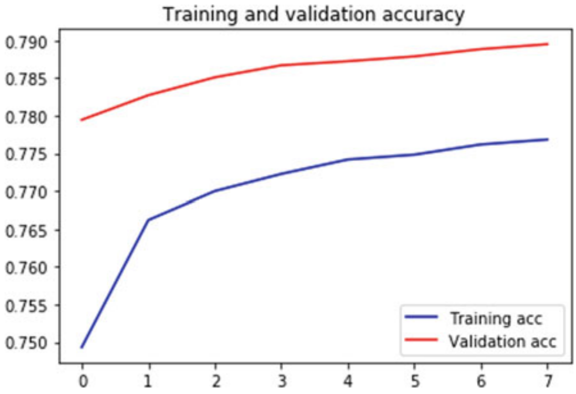


Fig. 6 Training and validation accuracy graph

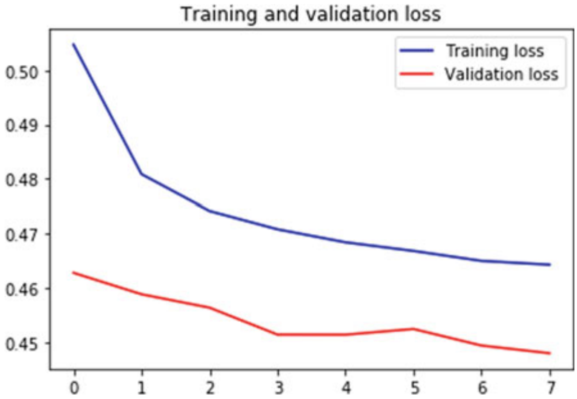


Fig. 7 Graph for training and validation loss

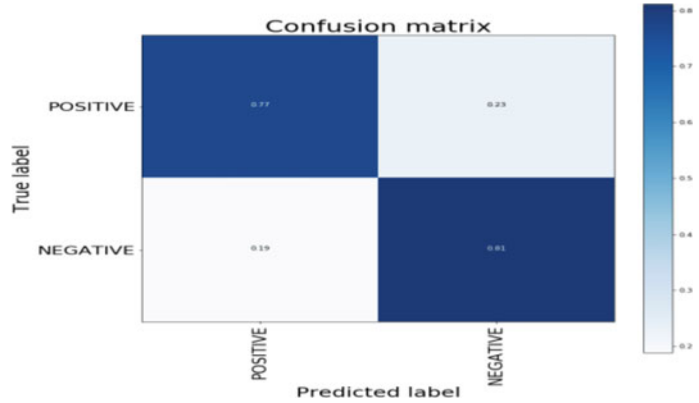
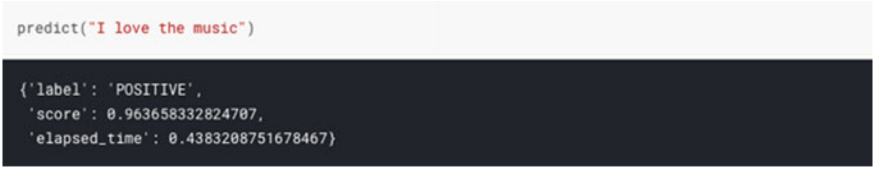


Fig. 8 Confusion matrix

This shows that our model is predicting scores and classifying sentences containing multiple slangs with very fewer errors. The proposed model is performing quite well-considering slangs in the dataset [16].

Now, let us predict some text and see how well out model is doing. Figures 9, 10, 11 and 12 show some examples of the score predicted by our model.

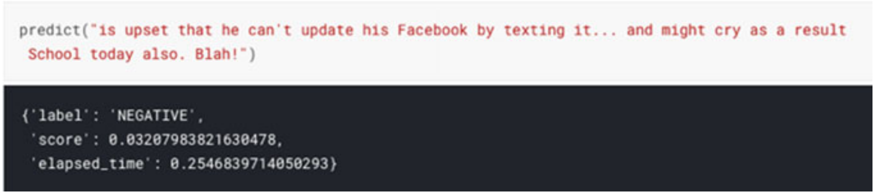
Hence, our model is giving scores from 0 to 1 up to 8 decimal places which can be further extended to rating inference from 0 to 10, where 0 being worst and ten being the best and can be applied on microblogs for analyzing and mining contents. Though our model performing great, there are still some future works that can be done like



```
predict("I love the music")
```

```
{'label': 'POSITIVE',
 'score': 0.963658332824707,
 'elapsed_time': 0.4383208751678467}
```

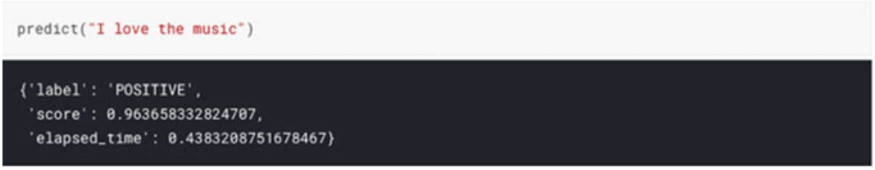
Fig. 9 Predicted score of "I love music"



```
predict("is upset that he can't update his Facebook by texting it... and might cry as a result
 School today also. Blah!")
```

```
{'label': 'NEGATIVE',
 'score': 0.03207983821630478,
 'elapsed_time': 0.2546839714050293}
```

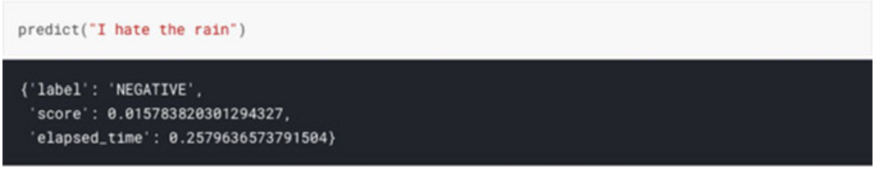
Fig. 10 Predicted score of "is upset that he can't update his Facebook by texting it … and might cry as a result School today also. Blah!"



```
predict("I love the music")
```

```
{'label': 'POSITIVE',
 'score': 0.963658332824707,
 'elapsed_time': 0.4383208751678467}
```

Fig. 11 Predicted score of "I love the music"



```
predict("I hate the rain")
```

```
{'label': 'NEGATIVE',
 'score': 0.015783820301294327,
 'elapsed_time': 0.2579636573791504}
```

Fig. 12 Predicted score of "I hate the rain"

Table 1 Comparison of the accuracies of various classifiers of the proposed method

S. No.	Name of the classifier	Accuracy (%)
1	LSTM network	78.99
2	Logistic regression	71
3	CNN	73

extending its from microblog to excellent contents like books, documentary, etc. and further improving it for multiple grams content.

Perform a sentimental analysis of Twitter slangs using LSTM Network, logistic regression, and CNN. Table 1 shows a comparative performance of the different techniques or classifiers used in the proposed method. Figure 13 shows accuracy and comparison of the LSTM network, logistic regression and CNN algorithms. Among these algorithms, the LSTM network gives the highest accuracy of 78.99% and it is shown in Fig. 13.The accuracy comparison of the proposed and existing methods is shown in Fig. 14.

5 Conclusion and Future Scope

The proposed work utilized Twitter data (sentiment140 dataset) to identify sentiment. After pre-processing the data, develop an algorithm that not only focuses on keywords but also takes into consideration every single word and tries to overcome slang by making a custom dataset and using stemmers to find its original keyword and meaning. Next, trained and tested the model by using various algorithms like long short-term memory network, logistic regression, and convolution neural network algorithms. Among these algorithms, the long short-term memory network gives the

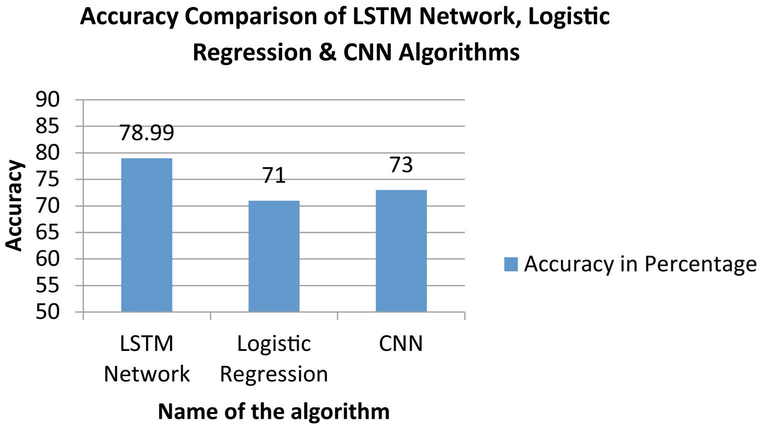


Fig. 13 Accuracy comparison of the LSTM network, logistic regression and CNN algorithms

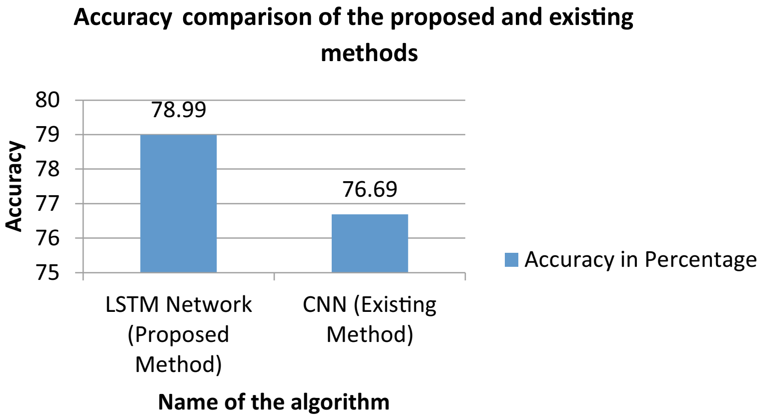


Fig. 14 Accuracy comparison of the proposed and existing methods

highest accuracy of 78.99%. In the future, to increase the accuracy of the proposed model, use bidirectional encoder representations from transformers and embeddings from language models algorithms for classification.

References

1. Dave K, Lawrence S, Pennock DM (2003) Mining the peanut gallery: opinion extraction and semantic classification of product reviews. In: WWW '03: proceedings of the 12th international conference on World Wide Web, pp 519–528

2. Pang B, Lee L, Vaithyanathan S (2002) Thumbs up? Sentiment classification using machine learning techniques. In: EMNLP ’02: Proceedings of the ACL-02 conference on empirical methods in natural language processing. Association for computational linguistics, pp 79–86
3. Turney PD (2002) Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: ACL ’02: Proceedings of the 40th annual meeting on association for computational linguistics, pp 417–424
4. Goldberg AB, Zhu X (2006) Seeing stars when there aren’t many stars: graph-based semi-supervised learning for sentiment categorization. In: Text Graphs ’06: proceedings of text graphs: the first workshop on graph based methods for natural language processing on the first workshop on graph based methods for natural language processing, association for computational linguistics, pp 45–52
5. Ki Leung CW, Fai Chan SC, Lai Chung F (2006) Integrating collaborative filtering and sentiment analysis: a rating inference approach. In: ECAI 2006 workshop on recommender systems, pp 62–66
6. Pang B, Lee L (2005) Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales. In: ACL’05: proceedings of the 43rd annual meeting on association for computational linguistics, pp 115–124
7. Gamon M (2004) Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis. In: COLING ’04: proceedings of the 20th international conference on computational linguistics, pp 841
8. Hu M, Liu B (2004) Mining and summarizing customer reviews. In: KDD ’04: proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pp 168–177
9. Manuel K, Varma K, Radha I, Krishna P (2010) Analyzing internet slang for sentiment mining. In: IEEE, 2010 second vaagdevi international conference on information technology for real world problems
10. Sun Z, Xu Y (2019) Slang detection and identification. In: Proceedings of the 23rd conference on computational natural language learning, pp 881–889
11. Mounika M, Srinivasa Gupta N, Valarmathi B (2020) Detection of depression related posts in tweets using classification methods—a comparative analysis. In: Proceeding of the international conference on computer networks, big data and IoT (ICCBI-2019), vol 49, pp 620–630
12. Srinivasa Gupta N, Valarmathi B, Sanju J (2012) Sentiment analysis using representative terms—a grouping approach for binary classification of documents. J Theor Appl Inf Technol 44:161–165
13. Valarmathi B, Srinivasa Gupta N, Palanisamy V (2016) Mahalanobis distance—the ultimate measure for sentiment analysis. Int Arab J Inf Technol 13:252–257
14. NhanCach D, María NM-G, Fernando De la P (2020) Sentiment analysis based on deep learning: a comparative study. Electron Open Access J 9:1–29
15. Go A, Bhayani R, Huang L (2009) Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford, vol 1
16. https://towardsdatascience.com/sentiment-analysis-of-transliterated-texts-in-hindi-and-marathi-languages-c29eab19fae2

Exploration of Varied Feature Descriptors for Diabetic Retinopathy Through Image Classification

K. Sreekumar and E. R. Vimina

Abstract Diabetic retinopathy (DR) is a disorder affecting retinal blood circulation; it is precipitated by diabetes mellitus. Since DR can lead to complete blindness, its early detection is of utmost importance. The computational methods for detecting DR include segmentation, feature extraction, and classification. This study uses different feature descriptors—local directional pattern (LDP), local binary pattern (LBP), and histogram-oriented gradients (HOG) for extracting features by processing diabetic retinopathy images. A set of supervised learning classifiers—support vector machine (SVM), random forest (RF), and K-nearest neighbor (KNN)—are used to classify the retina images based on the features extracted. The study indicates that a combination of LBP and HOG coupled with RF is the most accurate form among the various candidate approaches and gives the best accuracy of 87.50%.

Keywords Diabetic retinopathy · Retinal fundus image · Machine learning · Classification · Image processing · Feature extraction

1 Introduction

Diabetes is a chronic disease formed by an abnormal increase in blood sugar in the body. In 2015, as per estimates made by the International Diabetes Federation (IDF), there were 415 million diabetes affected persons in the world by 2040; the number is projected approximately to surge to 642 million that is 10% of all adult humans. Diabetic patients often face severe complications of diabetic retinopathy (DR) [1]. It affects the retina of the patients, and if left untreated, it can lead to blindness. A popular domain where computer vision is applied in object recognition or object classification the principal intent of this domain is extraction of features from the

---

K. Sreekumar (✉) · E. R. Vimina
Department of Computer Science and IT, Amrita Vishwa Vidyapeetham, Kochi, India
e-mail: sreekumar4@gmail.com

Amrita School of Arts and Sciences, Amrita Vishwa Vidyapeetham, Kochi, India

E. R. Vimina
e-mail: vimina.er@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_34

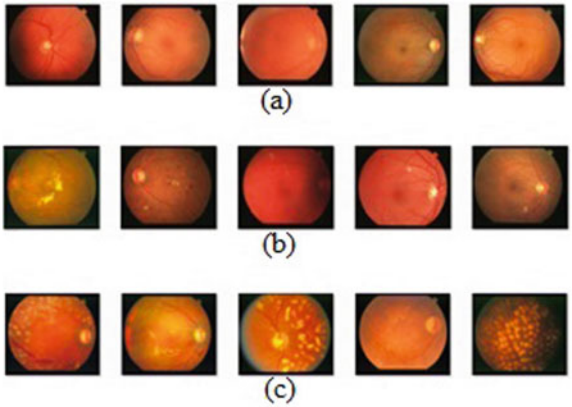


Fig. 1 **a** Normal, **b** non-proliferative diabetic retinopathy, **c** proliferative diabetic retinopathy

images and classification of images into suitably defined classes for this, any one of several possible classifiers or methods of classification could be used [2]. The various classification methods used here are SVM, RF, and KNN. A STARE dataset, containing around 400 images related to eye diseases, is used for the analysis. Around 131 retina images from the above set belonging to three categories, namely proliferative diabetic retinopathy (PDR), non-proliferative diabetic retinopathy (NPDR), and normal are selected interrelated to the diabetic retinopathy detection. Figure 1 shows a sample of retinal fundus images that are taken from the STARE dataset. The sample images on the first row belong to (a) normal category, the second row contains images showing (b) non-proliferative diabetic retinopathy, and the last row contains the images which show (c) proliferative diabetic retinopathy.

2 Related Work

Diabetic retinopathy can be detected by screening retinal fundus images of normal and diabetic patients. The optic disk, blood vessels, and fovea are the standard features of the fundus images. The blot hemorrhages and exudates are the main uncharacteristic features of diabetic retinopathy [3]. The features that mark NPDR are microaneurysms, dot and blot hemorrhages, flame-shaped hemorrhages, and retinal edema. The different features indicating PDR are new vessels at the disk (NVD) and new vessels elsewhere (NVE). The exudate can be detected by using strategies such as thresholding, edge detection, and classification. Backpropagation in a neural network can be used for feature selection to detect hemorrhage areas in the retinal fundus image.

A lot of work has already been done on methods for classification of DR and results have been largely encouraging. Pratt et al. proposed a convolutional neural network (CNN) approach, an area of deep learning for diagnosis of DR by analysis of digital fundus images as well as quantification of its severity [4]. They set up a network based on CNN architecture and incorporating data intensification and this resulted in a sensitivity of 95% from among 5000 images for validation. The accuracy was 75%.

Bhatia et al. presented a paper that seeks to arrive at decisions about the presence or absence of the disease by employing a collection of machine learning and classification algorithms [5]. Chandran et al. proposed an algorithm for spot wise extraction of features from the retinal shallow [6]. A random forest classification procedure was carried out on the features resulting from image spots. The approach displays enhanced specificity in comparison to previously existing methods and is suitable for deployment as a tool that performs screening for diabetic retinopathy.

Bui et al. described a method of automated dissection and used it to identify cotton wool patches on retina images as a marker of diabetic retinopathy [7]. Identification of cotton wool spots could help in the prevention of serious damage which could even lead to a total loss of vision. Their study demonstrates that cotton wool could be segmented with the accuracy was 85.54%; sensitivity, 85.9% and specificity, 84.4%. Carrera et al. proposed a diagnostic technique that uses digital processing performed on retinal images to achieve early detection of diabetic retinopathy. The objective is to achieve an automatic classification in terms of non-proliferative diabetic retinopathy (NPDR), grade found in a given retinal image [8]. This proposal has been tested and achieved 94% predictive capacity and a maximum sensitivity of 95%.

Arora et al. proposed a solution to overcome the problem generated using a machine learning method employing techniques of deep learning along with CNN automatically does pattern identification and achieves classification of imageries of retina among 5 classes [9]. The model's learning rate was 0.001 resulting in 74% validation accuracy. Alzami et al. proposed a method that works with the MESSIDOR dataset and employs a random forest classification and found that fractal dimensions are capable of distinguishing between healthy persons and those affected with diabetes retinopathy [10].

Bindhu proposed a method employing the semantic segmentation model for the biomedical images reducing the diagnostic time [11]. Samuel proposed a smart algorithm used for image processing which will be beneficial for the visually challenged by utilizing the text recognition, extraction of vocalization, and information [12]. Abraham proposed a method of assessment based on medical imaging technology (IoT-based) for the application used in health care [13].

Different types of techniques, architectures, simulations, and frameworks introduced by many researchers have played an important role in detecting lesions in the early stage of DR. Extraction of features of retinal fundus image and application of machine learning techniques to them is one of the methods to detect DR. Another method for classification of retinal fundus images, deep learning can be used in the field of DR. Deep learning is a class of machine learning algorithms that uses multiple layers to gradually excerpt higher-level features from the raw input. The convolutional neural network (CNN) approach, an area of deep learning, is an advanced method used for image classification to detect DR.

Thus, it is found that the various feature descriptors play a major role in extracting the required features and help the classifiers to perform the classification on retinal fundus images. In this study, diverse feature extraction methods like LBP, LDP, and HOG along with analysis of the various classification methods like SVM, KNN, and random forest are done on a set of retina images and the net accuracy for each combination is observed to select the best one.

3 Proposed Work

A comparative analysis is performed for classifying a set of diabetes retinopathy images by using different feature descriptors like LBP, LDP, and HOG along with machine learning algorithms. The classifiers like SVM, KNN and random forest will quantify the net accuracy of the experiment. A combination of LDP + HOG and LBP + HOG is also tested with the above mentioned classifiers to find the best accuracy rate.

The retina images are taken from a STARE dataset, which contains around 400 images and in that 131 images are related to DR. After the image acquisition, the same are preprocessed by applying green channel extraction to avoid noisy data. Various feature descriptors are applied to perform the feature extraction followed by classification done by the classification algorithms to obtain the various category labels, and finally, the net accuracy is retrieved. Figure 2 depicts the flow chart for the proposed work.

In this experiment, retina images which belong to three different classes namely proliferative diabetic retinopathy (PDR), non-proliferative iabetic retinopathy (NPDR), and normal are used. The feature descriptors LBP, LDP, and HOG are used for feature extraction along with machine learning classification algorithms SVM, RF, and KNN to retrieve the accuracy rate. 70% of the total 131 dataset images are trained in which the preprocessing and the feature extraction are executed. The extracted features are stored in a feature database. The remaining 30% of the dataset images are tested and classified into three categories.

3.1 Preprocessing

The preprocessing is done on the retina images collected from the STARE dataset. The retinal fundus images are RGB color images, which consist of three channels such as red, green, and blue. The green channel is alone extracted from this concatenated RGB, since the green channel will have less noise compared to the other two channels.

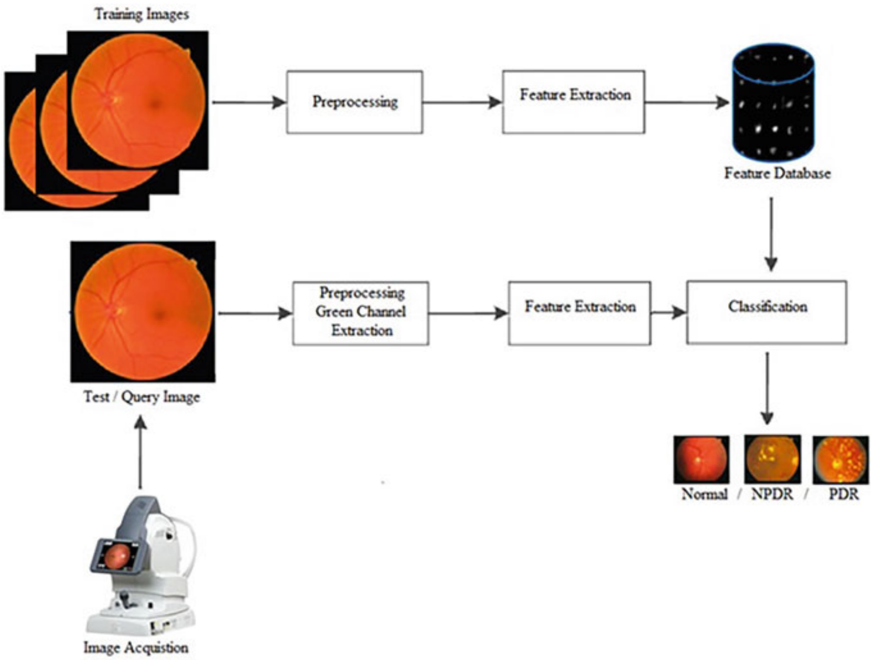


Fig. 2 Flow chart

3.2 Feature Extraction

Feature extraction comprises revealing and segregation of features of a given image such as shape, shading, baseline—and these features are combined to create a vector. The extraction of features is most crucial, as the specific features made available for selection determine the effectiveness of classification.

3.3 Local Binary Pattern (LBP)

In computer vision, the LBP, a visual descriptor, is used for classification. It marks image pixels by relating a threshold to the neighborhood of the pixel and extracts a specific binary number as result. This texture operator has gained acceptance in diverse applications, primarily due to its power of discrimination and essential simplicity [14].

Local binary pattern (LBP) is a non-parametric descriptor; it aims to summarize with efficiency, local structures in images. LBP works among adjacent pixels by fixing a center threshold. Checking if the neighboring value of the pixel exceeds or equals that of the pixel selected as the center which indicates it as 1, else 0. The decimal equivalent of the 8-digit binary number obtained is set to the central-pixel threshold. By applying a threshold to every pixel, obtain an LBP code corresponding to each pixel value.

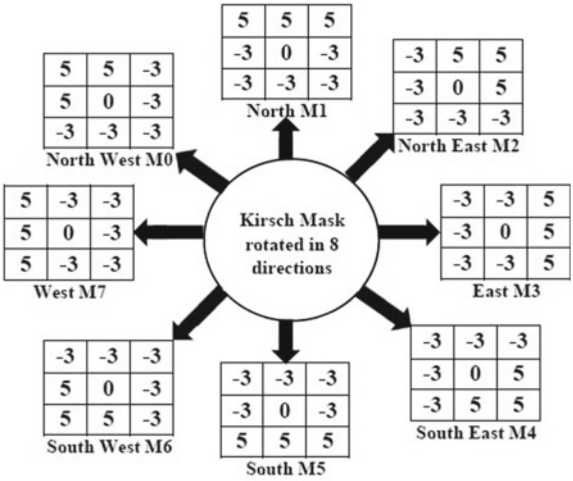
3.4 Local Directional Pattern (LDP)

The various challenges of the LBP such as noise and illumination change can be overcome easily by using the feature extraction method local directional patterns (LDP). It is an image feature that calculates the address values corresponding to the strings in various directions and then applies these values to work out the encoding of image textures [15, 16]. It is a code (8 bit) assigned to each pixel constituting the picture and the same is worked out in multiple methods.

Kirsch masks m_i—a typical edge detector, processes the mask values, with i ranging among values 0–7, founded on their position concerning 8 possible orientations, given a central-pixel within the picture. Figure 3 shows those masks. Along with some directions, the intervention of an edge or corner is indicative of greater reaction values is to understand the k's almost all significant routes for the above descriptor. The b_i is assigned to the highest retort shown by the guiding bit, k. All the other bits are set to 0. Finally, it jumps to Eq. (1). m_k indicates the kth ranked among relevant directional responses. The m_i is the eight directional edge responses and b_i is the bit response.

$$\mathrm{LDP}_K = \sum_{i=0}^{7} b_i(m_i - m_k) \times 2^i \tag{1}$$

Fig. 3 Kirsch masks



$$b_i(a) = \begin{cases} 1 & a \geq 0 \\ 0 & a < 0 \end{cases}$$

3.5 Histogram-Oriented Gradients (HOG)

Image appearance and shape can be denoted with HOG, which is a simple and efficient technique for extracting features. The shape and appearance of an image are denoted with HOG; the image is divided into tiny cells and edge directions calculated. To enhance precision, normalize the histograms. The HOG descriptor is centered on the design or form of the objects and the decision is made whether the pixel is part of an edge or not. The edge directions given by HOG are found in confined segments. It splits the full image into constituent gradients and parts; then directions corresponding to every region are calculated. The HOG yields a unique histogram for each of these areas and these histograms are formed out of pixel value gradients and orientations [17].

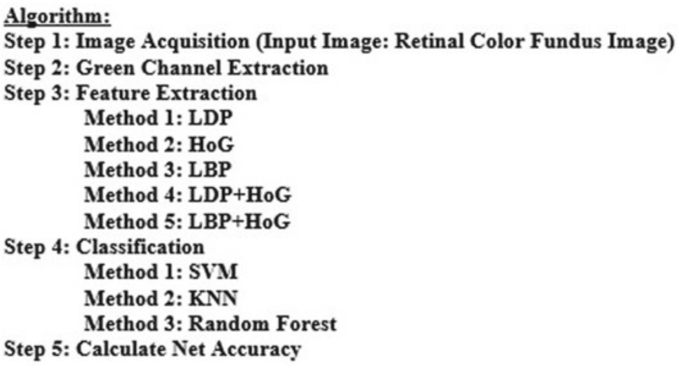
Each tiny patch is chosen from the images and then the gradients are worked out. The pixel values are needed for the patch. To estimate the x-axis alteration (G_x), subtract the pixel value to the left from the rightward pixel value. Next, subtract the beneath pixel value from the overhead pixel value of the selected pixel to find the y-direction gradient (G_y). The same cycle is repeated for every pixel in the image. Now, display the magnitude as well as the direction of every value of a pixel with the help of the x and y gradients computed in the prior stage. The magnitude, M—overall gradient vector for a point (x, y) is obtained from its Euclidean vector norm in Eq. (2).

$$M = \sqrt{\left(G_x^2 + G_y^2\right)} \tag{2}$$

The gradient vectors, direction, or angle (α) for a point (x, y) can be computed using Eq. (3).

$$\alpha = \tan^{-1}\left(G_y/G_x\right) \tag{3}$$

The histograms yielded by the HOG function descriptor for the full image are not established instead it subdivide the picture into 8 × 8 cells and for every cell determine the histogram of each directed gradient. Thus, for smaller patches, the histogram (or features) represents the full image. This value here certainly gets altered to 16 × 16 or 32 × 32 from 8 × 8 and the maximum benefit occurs in the path of the bin with respect to the pixels' location.



```
Algorithm:
Step 1: Image Acquisition (Input Image: Retinal Color Fundus Image)
Step 2: Green Channel Extraction
Step 3: Feature Extraction
        Method 1: LDP
        Method 2: HoG
        Method 3: LBP
        Method 4: LDP+HoG
        Method 5: LBP+HoG
Step 4: Classification
        Method 1: SVM
        Method 2: KNN
        Method 3: Random Forest
Step 5: Calculate Net Accuracy
```

Fig. 4 Algorithm

3.6 *Classification Methods*

The classifier algorithms like SVM, RF, and KNN are used for classifying various retina images. Support vector machine (SVM) was introduced by Vapniket et al. [5, 18]. Training lengthy decision trees with several train patches are the main feature of the random forest algorithm. [5, 6]. KNN classifier works according to the distance/similarity function that classifies based on similarity measure [19].

Figure 4 explains the algorithm for the proposed work. It begins with image acquisition, followed by green channel extraction and then the feature extraction using LDP, HOG, LBP, LDP + HOG, and LBP + HOG, and finally the classification using the machine learning algorithms SVM, KNN, and random forest.

4 Result Analysis

4.1 *Data Set and Implementation*

A STARE dataset is used for this analysis. STARE stands for structured analysis of the retina; the University of California initiated this project. From around 400 images, 131 images belong to three different classes, namely PDR, NPDR, and Normal. Table 1, shows the total number of images in each category. The analysis is done on every

Table 1 Category of images

Category	No. of images
Normal	40
Proliferative diabetic retinopathy (PDR)	22
Non-proliferative diabetic retinopathy (NPDR)	69

image and they are grouped among the above categories based on the specifications. The computational algorithms for the classification of retina images using machine learning techniques have been formulated and implemented using MATLAB 2018a. A GUI interface has been created for selecting the corresponding feature descriptor and the classifier combination to find the net accuracy rate.

A classifier evaluation is done based on the true positive (TN), true negative (TN), false positive (FP), and false negative (FN), which are the values obtained based on a two-class domain. The TP is the positive tuples and the TN is the negative tuples correctly labeled by the classifier. The FP is the negative tuples incorrectly labeled as positive and FN is the positive tuples incorrectly labeled as negative by the classifier. The accuracy measure is computed using Eq. (4) as shown below.

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

The entire dataset is divided into a training set (70%) and a testing set (30%) and the net accuracy is computed. In our experiment, the net accuracy of different feature descriptors like LBP, LDP, and HOG are coupled along with a set of classifiers—SVM, RF, and KNN are compared. Our study clearly shows that the combination of LBP and HOG coupled with RF gives the maximum accuracy (87.50%) among the various approaches. The observations shown in Table 2 describe experimental outcomes for the comparison of the net accuracy percentage for various classification methods and feature extraction combinations used. It is found that the combination of LBP + HOG along with the random forest algorithm is giving the highest accuracy of 87.50%.

The SVM classifier is combined with the different feature descriptors LDP, HOG, LBP, and the net accuracy of the respective combination is shown in the above table. The SVM is also used to calculate the accuracy rate by combining LBP + HOG and LDP + HOG and the application result is also shown in the above table. Similarly, the KNN classifier is combined with LDP, HOG, and LBP to get the respective accuracy rate as shown. The KNN is also applied with the combination of LBP + HOG and LDP + HOG to get the accuracy rate as shown in the above table. Finally, the random forest classifier is used along with LDP, HOG, and LBP to get the results. It is also used with the combination of LBP + HOG and LDP + HOG and the respective results are recorded in the above table.

Table 2 Comparison of net accuracy rate

	SVM (%)	KNN (%)	RF (%)
LDP	50.00	47.50	47.50
HOG	52.50	50.00	55.00
LBP	52.50	57.50	82.50
LDP + HOG	50.00	47.50	52.50
LBP + HOG	52.50	45.00	87.50

4.2 Discussions

The main advantages of SVM are that it works well with a clear margin of separation, memory efficient, and is very effective in high dimensional spaces. It is also suited for extreme case binary classification. The disadvantages of SVM are a higher training time needed for large datasets, unsatisfactory performance with noisy data, and that it will not directly provide probability estimates. An advantage of the KNN algorithm is the ease of understanding and implementation. It tags the new entry data based on the history available. It quickly responds to the changes in the input during real-time use and it automatically adjusts to multi-class without extra effort. The disadvantage of the KNN algorithm is that despite its simplicity, its efficiency and speed degrade for large data sets. KNN is very sensitive to outliers since it chooses the neighbors using the distance criteria. Random forest algorithms are a group of decision trees for prediction; they take input from all the trees and gives the prediction. This reduces the overall variance and error and thus improves accuracy. The other advantages of the random forest algorithm include the capability to handle huge amounts of data, to handle missing data, and robustness to outliers it is very useful for feature extraction. It has the following disadvantages: the predictions of the trees need to be uncorrelated and the features should have some predictive power else they give some error. The complexity and longer training period are some other minus points of the random forest algorithm.

Figure 5 depicts the net accuracy rate of classification approaches for a combination of feature descriptors. The chart clearly indicates the net accuracy rate for LDP, HOG, LBP, LDP + HOG, and LBP + HOG combined with the three classification algorithms SVM, KNN, and random forest. The combination of LBP + HOG along with the random forest algorithm gives the highest accuracy of 87.50%.

70% of the dataset has been used for the training of the model and the remaining 30% of the dataset, for testing. The implementation of the above comparative study is done using MATLAB 2018a. A graphic user interface has been created using

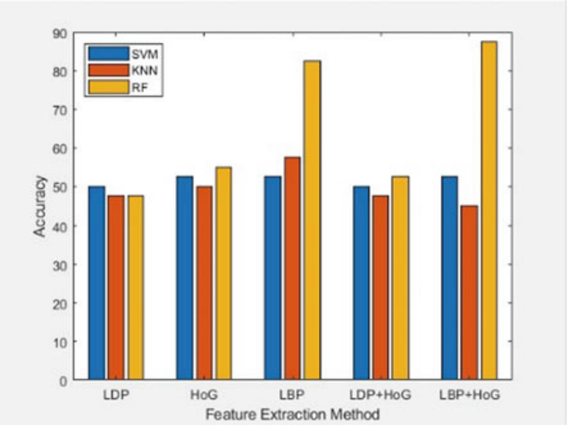


Fig. 5 Net accuracy rate of classification approaches for combination of feature descriptors

MATLAB to enable the selection of various simulation parameters. It is possible to select the classifier along with the feature extractor combination with the help of controls placed in the graphic user interface. The interface enables the display of the accuracy of the selected combination by a button click and there is also an option to print the net accuracy rate graph.

5 Conclusion

This article aimed to formulate an efficient feature extraction method for retina images by coupling different feature extraction methods with various machine learning techniques. The process was tested using a structured analysis of the retina dataset which contains 131 images, further grouped into 3 classes. The experiment results show that the combination of local binary pattern and histogram-oriented gradients feature extraction methods coupled with the random forest algorithm gives the best accuracy (87.50%), which compares favorably with the other combinations tested. It is suggested that this combination could be used in the analysis of retina images in particular to achieve accurate and timely classification that enables ophthalmologists to prioritize cases and identify emergencies that need immediate intervention.

References

1. Marina T et al (2002)Perceptions of diabetic retinopathy and screening procedures among diabetic people. Diab Med 19(10):810–813
2. Carmen V et al (2016)Automated detection of diabetic retinopathy in retinal images. Indian J Ophthalmol 64(1):26
3. Dilip Singh S, Nair S, Khobragade P (2017)Diabetic retinal fundus images: preprocessing and feature extraction for early detection of diabetic retinopathy. Biomed Pharmacol J 10(2):615–626
4. Harry P et al (2016)Convolutional neural networks for diabetic retinopathy. Procedia Comput Sci 90:200–205
5. Karan B, Arora S, Tomar R (2016)Diagnosis of diabetic retinopathy using machine learning classification algorithm. In: 2016 2nd international conference on next generation computing technologies (NGCT). IEEE
6. Anaswara C, Nisha KK, Vineetha S (2016)Computer aided approach for proliferative diabetic retinopathy detection in color retinal images. In: 2016 international conference on next generation intelligent systems (ICNGIS). IEEE
7. Toan B, Maneerat N, Watchareeruetai U (2017)Detection of cotton wool for diabetic retinopathy analysis using neural network. In: 2017 IEEE 10th international workshop on computational intelligence and applications (IWCIA). IEEE
8. Carrera EV, González A, Carrera R (2017)Automated detection of diabetic retinopathy using SVM. In: 2017 IEEE XXIV international conference on electronics, electrical engineering and computing (INTERCON). IEEE

9. Mamta A, Pandey M (2019)Deep neural network for diabetic retinopathy detection. In: 2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon). IEEE
10. Farrikh A, Megantara RA, Fanani AZ (2019)Diabetic retinopathy grade classification based on fractal analysis and random forest. In: 2019 international seminar on application for technology of information and communication (iSemantic). IEEE
11. Bindhu V (2019) Biomedical image analysis using semantic segmentation. J Innov Image Process (JIIP) 1(02):91–101
12. Manoharan S (2019) A smart image processing algorithm for text recognition information extraction and vocalization for the visually challenged. J Innov Image Process (JIIP) 1(01):31–38
13. Chandy A (2019) A review on iot based medical imaging technology for healthcare applications. J Innov Image Process (JIIP) 1(01):51–60
14. Prithaj B et al (2018)Local neighborhood intensity pattern–a new texture feature descriptor for image retrieval. Exp Syst Appl 113:100–115
15. Zhou J, Tianwei Xu, Gan J (2013) Feature extraction based on local directional pattern with svm decision-level fusion for facial expression recognition. Int J Bio-Sci Bio-Technol 5(2):101–110
16. Taskeed J, Kabir MH, Chae O (2010)Local directional pattern (LDP)–a robust image descriptor for object recognition. In: 2010 7th IEEE international conference on advanced video and signal based surveillance. IEEE
17. Navneet D, Triggs B (2005)Histograms of oriented gradients for human detection. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), vol 1. IEEE
18. Vladimir V (1998)The support vector method of function estimation. In: Nonlinear modeling. Springer, Boston, MA, pp 55–85
19. Wu X, Kumar V (eds) (2009) The top ten algorithms in data mining. CRC Press

News Topic Classification Using Machine Learning Techniques

Pramod Sunagar, Anita Kanavalli, Sushmitha S. Nayak, Shriya Raj Mahan, Saurabh Prasad, and Shiv Prasad

Abstract News topic classification is a method of classifying news articles available in text data into some predefined classes or labels. This is one of the applications of text classification. Text classification can be applied in the fields of spam filtering, language recognition, segmenting customer feedbacks, segregating technical documents, etc. This paper discusses news topic classification on AG's News Topic Classification Dataset using machine learning algorithms such as linear support vector machine, multinomial Naive Bayesian classifier, K-Nearest Neighbor, Rocchio, bagging, and boosting. This paper discusses three steps for classification, namely pre-processing of text, then applying feature extraction techniques, and finally implementing machine learning algorithms. These algorithms are compared using evaluation metrics like Accuracy, Recall, Precision, and F1 Score.

Keywords Text Classification · Natural language processing (NLPs) · Term frequency–inverse document Frequency (TF-IDF) · Support vector machine (SVM) · K-nearest neighbours (KNN) · Naïve Bayes · Rocchio

---

P. Sunagar (✉) · A. Kanavalli · S. S. Nayak · S. R. Mahan · S. Prasad · S. Prasad
Department of Computer Science & Engineering, Ramaiah Institute of Technology, Bangalore 560054, India
e-mail: pramods@msrit.edu

Visvesvaraya Technological University, Belagavi, Karnataka, India

A. Kanavalli
e-mail: anithak@msrit.edu

S. S. Nayak
e-mail: nayaksushmitha90@gmail.com

S. R. Mahan
e-mail: shriyarajmahan@gmail.com

S. Prasad
e-mail: saurabhprasad12@gmail.com

S. Prasad
e-mail: shivpsmy227721@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_35

1 Introduction

Text classification or text labeling is the method of categorizing text into prearranged groups [1]. Due to the recent breakthroughs in Natural Language Processing (NLP) and text mining, the application of text classification has reached many real-time applications. The machine learning algorithms can be implemented to automatically evaluate text and then allocate a set of pre-defined tags or categories based on its content. Due to the digital era and social media influence, the unstructured text is in large volumes. It is difficult to analyze such a large amount of text data if not organized. Unstructured textual data is nothing but the text which is available in its natural free form. Machines try to analyze this textual data by implementing pre-processing techniques and extract the information by implementing specific algorithms. Using NLPs, it has proven to be fast, cost-effective, and scalable. Text classifications are widely used in real-time applications like spam filtering, sentiment analysis, automated responses to customers, fraud detection, etc. [2, 3]. Today, the amount of social media data is increasing day by day as more and more people have started using the social sites and data is in great abundance, so this data is of no use until and unless it is made in some useful way of being classified under some categories which later can be studied on, this is the classifications that the text classifier provides us with. In this work, the AG's News Topic Classification Dataset was considered which had four classes, namely Business, Sci/Tech, Sports, and World. Text classification works like this, the training set $N = (N1,\ldots, Nn)$ consists of news articles that are already labeled with a class Business, Sci/Tech, Sports, and World. The task is then to establish a classification model which will assign the exact class to a new news article.

2 Literature Survey

Text classification or text categorization is the process of categorizing text into organized categories. By using NLPs techniques, text classifiers can automatically analyze text and then assign a set of previously defined tags or classes. During recent years, studies are ongoing on the classification of unstructured texts. Papers were published regarding text classification to classify papers using text mining which classified text using keywords, expectation maximization, and hierarchical shrinkage [4]. There were also studies to classify text using bootstrapping methods with keywords. This naive Bayesian classifier works by first selecting a class of document, and then document words are to be generated based on class-specific traits. After this, make use of the Expectation–Maximization (EM) algorithm for the entire dataset to give weighted class labels for all sets of document. This model can improve the model iteratively by first determining the missing values and then using the found values to improve the model [5]. Also papers on some text classification techniques, using WordNet Hypernyms. In this classification technique, different binary classification problems of changing difficulty are to be defined, and discrimination rules are produced for each problem so defined using ripper system and hypernym density representation. Rules can also be given without any use of WordNet by using the knowledge of recently used words. The last problem faced after doing the above mentioned steps is whether the improvements made by the process above can generalize some machine learning algorithms or not. The changes defined above in the process will grant some semantic character. More revealing ambiguity for words can provide us with more accurate features for hyponym [6]. Text categorization is the process of categorizing or classifying text in different groups based on keywords or phrases and then giving the categories that are already previously defined. There are many applications of this such as content classification, handling loads of unstructured data from social media, the spam filtering in emails, etc. Here main keywords are used from the documents to classify and assign them some category of documents. Such keywords provide almost every information of the document's content. Such keywords can be called a summary of its content [7].

Later, another method for text classification was proposed which was using text segmentation which was based on a feature selection method that gives its utmost emphasis on the document structure. Then, features were also extracted from the topic of the paper or document rather than just the content of the document. Some authors proposed to extract features as short as possible such as from titles, summary, abstract, etc. Various methods of obtaining new features were then introduced [8]. The next method studied is for making an efficient text classifier was PDF text classification using information retrieval. In a PDF, the content of it is usually unstructured or semi-structured and often mixed which makes it challenging to classify using NLP algorithms. Information Extraction (IE) systems can help us in getting the useful information that are needed to classify but most of these systems are not built to work with the PDFs which is a very important source for information retrieval. In a systematic review development, information extraction is a time-consuming process and can also be prone to so many errors. The main vision behind this is to categorize the documents using these information extraction systems. Normal machine learning algorithms occupy us with less accurate results than multi-pass NLP algorithms. These algorithms can also improve the performance of information extraction systems [9]. The next method so used in the classification is using support vector machines. Support vector machines have great importance in real-world problem solving and shown us some significant results. Next is the study of a new method of classification with a support vector machine which is an active learning process. The SVM understands the instances that are going to be used in the next stage and can request for it significantly. For this method, as it is described as active learning, it makes use of a training set that is already categorized to train the model that is being created. For this process, version space can be used. The center of the mass of a version space can be approximately found by using Bayes point machines. It is always better to minimize the labeled instances because labeling them is pretty much time taking process and can also be a bit costly. Make use of active learning to determine the predefined class labels rather than choosing random labels from a training set. Now text classification plays a very important role as in the modern world where the amount of unstructured data is huge and there is a need for a system so it can categorize this data and have easy access to it [10].

3 Implementation Details

3.1 Pre-processing

(a) **Tokenization**: It is a mechanism where a text or string is split into a set of tokens [11, 12]. Large text documents can be split into sentences, sentences that can be split into words, etc.

(b) **Stop Word Removal**: Once the text is tokenized, the next step is to remove the stop words which help in differentiating words with specific meaning from the ones which add a specific structure to the sentence.

(c) **Stemming**: This technique helps in reducing a word to a stem or base form [13, 14]. Snowball stemmer is a popular stemmer for the English Language.

(d) **TF-IDF**: Term frequency–Inverse Document Frequency suggests how important or relevant a word is for a document in the group of documents. TF-IDF for a word from a document is calculated by multiplying term frequency and inverse document frequency metrics [15]. The term frequency metric will give you the count of instances of a word in the document. Inverse document frequency specifies how important a word is to a document if it is appearing only for one particular document or how common the word is to all the documents if it appears in all the documents.

This section briefly explains the design of the proposed method. State diagrams indeed consist of a finite number of states and sometimes show a reasonable abstraction. All states eventually lead to a final state where the class label of a text document being given in input is predicted and helps to categorize the text document to which label it should belong to. The training data is first evaluated and represented using feature vectors for further help in the classification process and eventually create a model that can certainly predict the classes for text documents. As in Fig. 1, the process of classification begins with data gathering or acquisition and feeding it to the text classification algorithm which then trains itself according to the data being fed to it. Based on the training data being given to the algorithm, it attains prediction accuracy and thus classifies a text under a certain category. More training data are being fed for better accuracy.

The new classification task is carried out by preprocessing the dataset first. For the process of processing the raw text, the snowball stemmer (a string-processing language) is used. In the pre-processing stage, the unwanted text is cleaned by removing exclamations, punctuations, semi-colons, etc. Then, the next step is to start filtering the data by removal of such data which does not add much useful information. The parameter and the feature extraction are based on the algorithms. In this

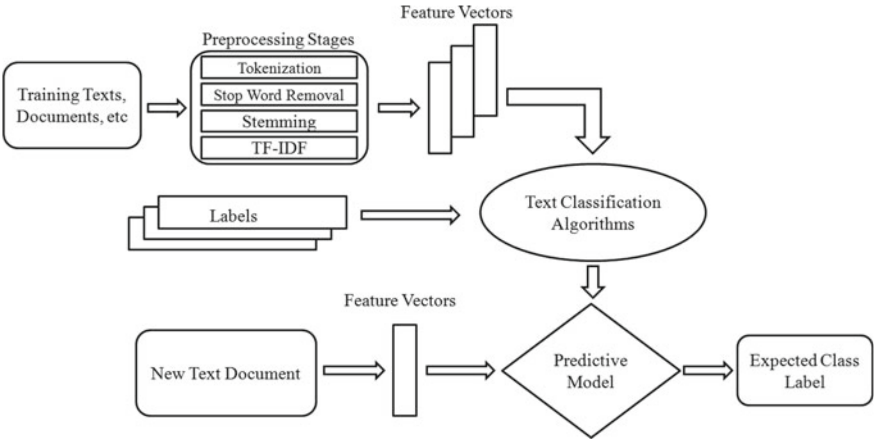


Fig. 1 Proposed architecture

work, a total of six machine learning algorithms are implemented and a comparative analysis is done on them using different evaluation metrics. The linear support vector machine (SVM) algorithm had shown the highest accuracy among the other algorithms.

The K-Nearest Neighbor (KNN) [16] algorithm classifies new data points depending on the similarity measures. The third algorithm which was implemented for the classification is the Rocchio algorithm [17]. This was developed using the vector space model and relevance feedback method. Then, bagging [18] and boosting [19, 20] algorithms were implemented for the classification. In bagging, the estimation uncertainty was reduced by producing additional data from the dataset to learn using variations of repetitions to generate multi-sets of the initial data. Boosting is an iterative procedure that fine-tunes the weight of an observation based on the final classification. Then, multinomial Naive-Bayes algorithm [21] was implemented which is based on the principle that every pair of features being classified is independent of each other. Finally, the linear SVM was implemented and found out that it is more accurate and has got better results of classification.

The support vector machine is a dominant and supervised learning algorithm commonly used for classification purposes. The SVM divides the classes by creating different hyperplanes. Then, the SVM [22] classifies the new samples using the hyperplanes. It was also considered how frequently a data was occurring or a word was occurring in the given text data. Only the important data was considered for the classification process. The important data is nothing but the data that came after the filtering stage and after the removal of useless and noisy data. Thus, a well-transformed data gave a good classification with the help of SVMs.

3.2 Dataset Used

The dataset used in this work was AG's News Topic Classification Dataset. This dataset consisted of four classes, namely Business, Sci/Tech, Sports, and World. There were 30,000 training samples and 1900 testing samples per class. The dataset consists of 120,000 training samples and 7600 testing samples. The files train.csv and test.csv contain all the training samples as comma-separated values. There are three columns in them, corresponding to class index (1–4), title, and description.

4 Evaluation Metrics

Initially, the pre-processing of the text was carried out, then feature extraction techniques were applied, and then the six machine learning algorithms were implemented on the test and training datasets. Now to figure out how effective these algorithms were while classifying, there was a need to assess them using a few evaluation metrics. For this purpose, the different metrics were used such as Accuracy, Recall, Precision, and F1 Score to exhibit the effectiveness of the algorithms.

(a) **Accuracy**: It is one of the simplest metrics. It is calculated as the number of appropriately predicted classes divided by the total number of predictions.

$$\text{Accuracy} = \frac{\text{NumberofCorrectPrediction}}{\text{TotalNumberofPredictionsMade}} \tag{1}$$

(b) **Recall**: It is another important metric and is calculated as the number of True Positives divided by the sum of True Positive and False Negative.

$$\text{Recall} = \frac{\text{TruePositive}}{(\text{TruePositives} + \text{FalseNegatives})} \tag{2}$$

(c) **Precision**: When the dataset has an imbalanced class distribution, then precision is the preferred metric compared to accuracy. It is calculated as the number of True Positives divided by the sum of True Positive and False Positive.

$$\text{Precision} = \frac{\text{TruePositives}}{(\text{TruePositives} + \text{FalsePositives})} \tag{3}$$

(d) **F1 Score**: This metric combines the Recall and Precision into one metric. It is the harmonic mean of recall and precision. It is defined as follows:

$$\text{F1Score} = 2 * \frac{1}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} \tag{4}$$

5 Results

5.1 Dataset Used

The performance of the machine learning algorithms is evaluated using four different evaluation metrics. The performance of all the algorithms is compared to the Accuracy, Recall, Precision, and F1 Score. The same is shown in Fig. 2 and in Table 1.

It was found out that the linear support vector machine performed better than all the other algorithms considered in this work. The support vector machine had consistently demonstrated 90% and above while predicting the correct labels. Multinomial Naïve Bayes comes next in the order when it comes to the prediction of correct labels. Boosting and Rocchio algorithms have demonstrated poor performance in predicting the true labels.

The training time and prediction time of all the algorithms are presented in Table 2. Boosting displayed the worst performance in terms of training time with 71.573469 s. While K-Nearest Neighbor algorithm is classified in very less time, i.e., 1.195145 s to train. Bagging displayed the worst performance in terms of prediction time with 157.886281 s. While the Naïve Bayes algorithm took very less time, i.e., 0.075812 s to predict.

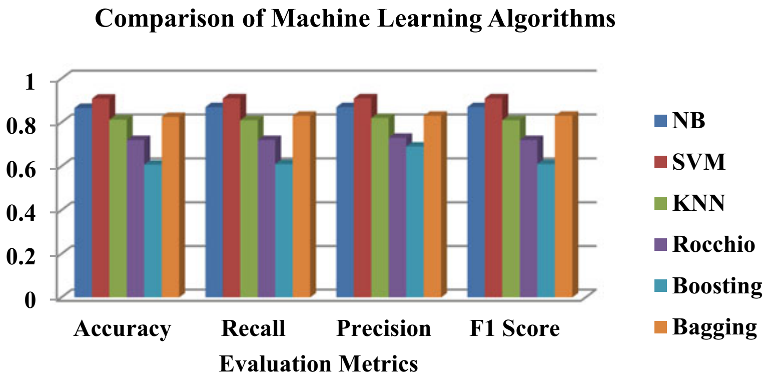


Fig. 2 Comparison of machine learning algorithms on various evaluation metrics

Table 1 Comparison of machine learning algorithms on different evaluation metrics

	NB	SVM	KNN	Rocchio	Boosting	Bagging
Accuracy	0.8662	**0.9085**	0.8124	0.7199	0.6075	0.8257
Recall	0.87	**0.91**	0.81	0.72	0.61	0.83
Precision	0.87	**0.91**	0.82	0.73	0.69	0.83
F1 Score	0.87	**0.91**	0.81	0.72	0.61	0.83

Table 2 Training time of different machine learning algorithms

	NB	SVM	KNN	Rocchio	Boosting	Bagging
Training time	1.482	10.974	1.195	1.226	71.573	1.520
Prediction time	0.075	0.586	16.167	0.0861	0.099	157.886

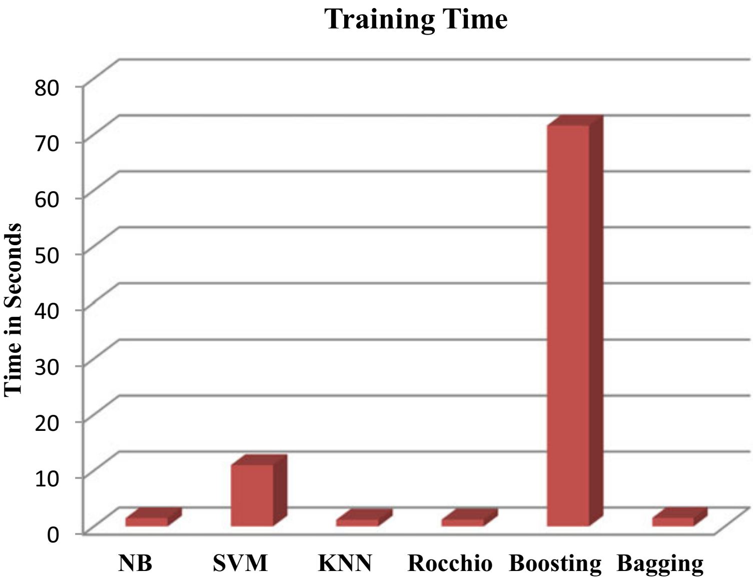


Fig. 3 Training time of different machine learning algorithms

In this work, the bar graphs are used for demonstrating the training time for all the algorithms considered. All the algorithms have demonstrated the short training time except for the boosting algorithm. The boosting algorithm is an iterative algorithm that keeps building strong learners. Due to this iterative nature of the boosting algorithm, it takes more time to train the model. The same has been shown in Fig. 3. The choice of algorithm for any real-time application depends on how fast the algorithm will predict the output. The more the prediction time, the less is the application of the algorithm in real-time applications. The prediction time for all the six machine learning algorithms is as shown in Fig. 4. The Naive Bayes algorithm took 0.075 s time to predict and bagging took 157.886 s to predict.

5.2 *Confusion Matrix*

The confusion matrixes are one more way to test the correctness of the model for any classification work. The matrix will give you the percentage of the correctly classified

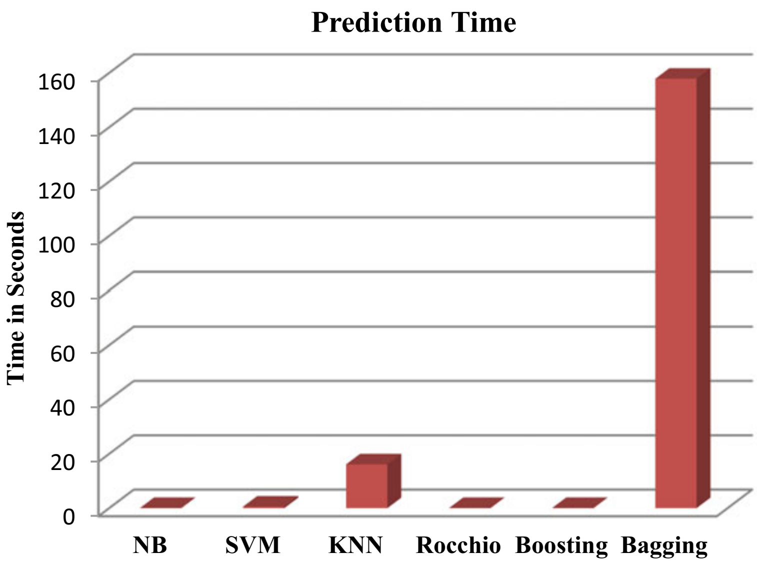


Fig. 4 Prediction time of different machine learning algorithms

data point. In the testing dataset, there are a total of 1900 rows per class. The Naive Bayes algorithm correctly predicts 1558 out of 1900 news topics with business labels. Similarly, the algorithm correctly predicts 1596, 1786, and 1653 news topics with Sci/Tech, Sports, and World labels, respectively. The same is shown as the confusion matrix in Fig. 5.

The linear support vector machine algorithm correctly predicts 1577 out of 1900 news topics with business labels. Similarly, the algorithm correctly predicts 1615, 1786, and 1672 news topics with Sci/Tech, Sports, and World labels, respectively. The same is shown in Fig. 6.

Fig. 5 Confusion matrix for multinomial Naïve Bayes

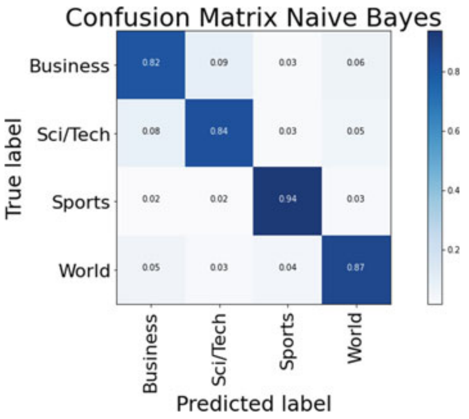
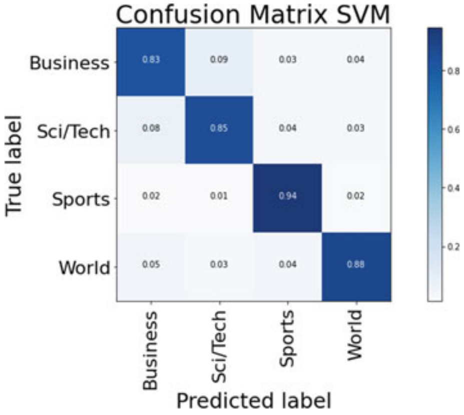


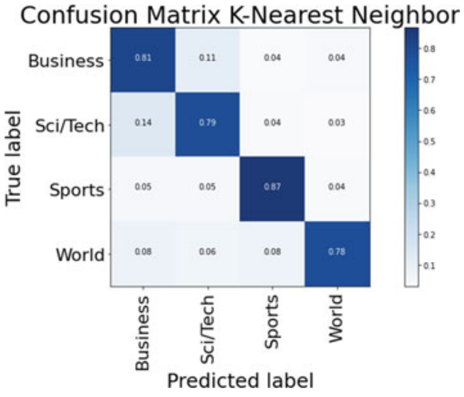
Fig. 6 Confusion matrix for linear support vector machine



The K-Nearest Neighbor algorithm correctly predicts 1539 out of 1900 news topics with business labels. Similarly, the algorithm correctly predicts 1501, 1653, and 1482 news topics with Sci/Tech, Sports, and World labels, respectively. The same is shown as the confusion matrix in Fig. 7. The Rocchio algorithm correctly predicts 1197 out of 1900 news topics with business labels. Compared to other algorithms, this algorithm has shown less accuracy due to the inconsistencies in the laws of classification between the various processes. Similarly, the algorithm correctly predicts 1387, 1596, and 1311 news topics with Sci/Tech, Sports, and World labels, respectively, and the same is shown in Fig. 8.

The Bagging algorithm correctly predicts 1501 out of 1900 news topics with business labels. The algorithm correctly predicts 1520, 1691, and 1539 news topics with Sci/Tech, Sports, and World labels, respectively, and the same is shown as the confusion matrix in Fig. 9. The boosting algorithm correctly predicts 893 out of 1900 news topics with business labels. The algorithm correctly predicts 969, 1672, and 1102 news topics with Sci/Tech, Sports, and World labels, respectively. This

Fig. 7 Confusion matrix for K-Nearest Neighbor



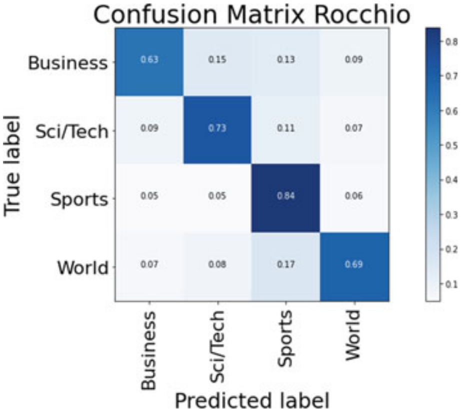


Fig. 8 Confusion matrix for Rocchio

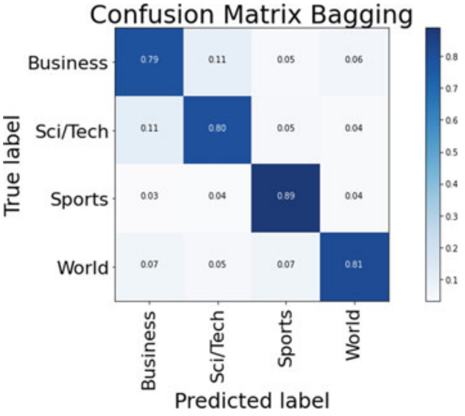


Fig. 9 Confusion matrix for bagging

algorithm displays the least accuracy compared to other algorithms, and the same is shown in Fig. 10.

5.3 Receiver Operating Characteristic (ROC)

A ROC curve or receiver operating characteristic curve is a graphical chart that indicates a binary classifier system's diagnostic capabilities as the threshold for discrimination is varied. Typically, the ROC curves are often used in binary classification to evaluate a classifier's performance. To extend the ROC curve and ROC region to multi-label classification, the output must be binarized. The ROC curve is formed by the plotting at different threshold settings of the True Positive rate (TPR) against the False Positive rate (FPR). There are four labels in the dataset, and the algorithms like Naive Bayes, SVM, KNN, bagging and boosting are implemented to generate ROC

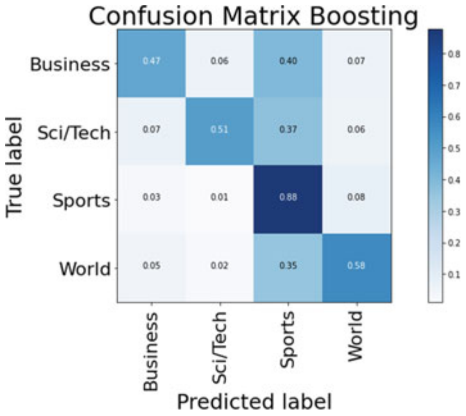


Fig. 10 Confusion matrix for boosting

for Business labels. To generate ROC for Rocchio, the algorithm has to compute a score. But by using the nearest centroid classifier, the algorithm displays the limitation to compute the score. Hence, the ROC curve for the Rocchio is not shown in Fig. 11.

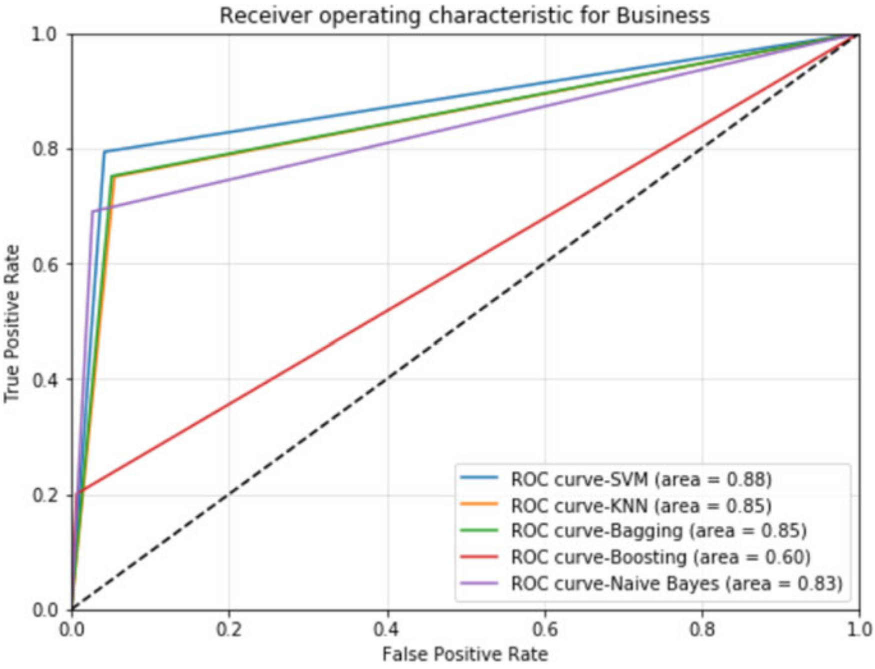


Fig. 11 ROC curve for machine learning algorithms for business label

6 Conclusion

The classification of News Topic was implemented using AG's News Topic dataset under predefined class labels. An approach was proposed to classify news texts comprising three steps. Text preprocessing was carried out first with a Snowball stemmer. The second step involved feature extraction based on the count vectorizer and TF-IDF vectorizer. Classification based on SVM, MNB, Rocchio, K-Nearest Neighbor, boosting, and bagging algorithms was the final step. For comparative analysis of these machine learning algorithms, different evaluation metrics like Accuracy, Precision, Recall, and F1 Score were used. The F1 Score was used to obtain a single value between precision and recall. As per the results, the SVM showed the highest accuracy of 91% among the algorithms considered, also requiring low training time and average prediction time. Multinomial Naïve Bayes showed the second-best accuracy among the algorithms, also requiring average training time and low prediction time. If the time requirements are the main focus, then the KNN would be efficient considering that the algorithm will deal with datasets of higher size. Linear SVM works poorly when the dataset is quite large and has more noise. The Naive Bayes classifier presumes highly about the structure of the data distribution. KNN is limited by data storage barriers for exploring nearest neighbors for large search problems. Boosting and bagging approaches have weaknesses such as the numerical complexity and lack of interpretability, which indicate that these models do not decide the function value. The Rocchio algorithm exhibits poor classification precision, and the linear combination is too uncomplicated for classification.

Acknowledgements This work was supported by Ramaiah Institute of Technology, Bangalore-560054, and Visvesvaraya Technological University, Jnana Sangama, Belagavi -590018.

References

1. Kowsari K, JafariMeimandi K, Heidarysafa M, Mendu S, Barnes L, Brown D (2019) Text classification algorithms: a survey. Information 10(4):150
2. Aggarwal CC, Zhai C (2012) A survey of text classification algorithms. In: Mining text data. Springer, Berlin/Heidelberg, Germany, pp 163–222
3. Aggarwal CC, Zhai CX (2012) Mining text data. Springer, Berlin/Heidelberg, Germany
4. Sulova S, Todoranova L, Penchev B, Nacheva R (2017) Using text mining to classify research papers. Int Multidisc Sci GeoConf Surv Geol Min Ecol Manag SGEM 17(21):647–654
5. McCallum A, Nigam K (1999) Text classification by bootstrapping with keywords, EM and shrinkage. In: Unsupervised learning in natural language processing
6. Scott S, Matwin S (1998) Text classification using WordNet hypernyms. In: Usage of WordNet in natural language processing systems
7. Menaka (2014) Text classification using keyword extraction technique
8. Nguyen TH, Shirai K (2013) Text classification of technical papers based on text segmentation. In: International conference on application of natural language to information systems. Springer, Berlin, Heidelberg, pp 278–284

9. Bui DDA, Del Fiol G, Jonnalagadda S (2016) PDF text classification to leverage information extraction from publication reports. J Biomed Inform 61:141–148
10. Tong S, Koller D (2001) Support vector machine active learning with applications to text classification. J Mach Learn Res 2:45–66
11. Verma T, Renu R, Gaur D (2014) Tokenization and filtering process in RapidMiner. Int J Appl Inf Syst 7(2):16–18
12. Aggarwal CC (2018) Machine learning for text. Springer International Publishing, Cham
13. Spirovski K, Stevanoska E, Kulakov A, Popeska Z, Velinov G (2018) Comparison of different model's performances in task of document classification. In: Proceedings of the 8th international conference on web intelligence, mining and semantics, pp 1–12
14. Singh J, Gupta V (2016) Text stemming: approaches, applications, and challenges. ACM Comput Surv (CSUR) 49(3):1–46
15. Jones KS (1972) A statistical interpretation of term specificity and its application in retrieval. J Doc
16. Jiang S, Pang G, Wu M, Kuang L (2012) An improved K-nearest-neighbor algorithm for text categorization. Expert Syst Appl 39(1):1503–1509
17. Korde V, Mahender CN (2012) Text classification and classifiers: a survey. Int J Artif Intell Appl 3(2):85
18. Breiman L (1996) Bagging predictors. Mach Learn 24(2):123–140
19. Freund Y (1992) An improved boosting algorithm and its implications on learning complexity. In: Proceedings of the fifth annual workshop on computational learning theory, pp 391–398
20. Bloehdorn S, Hotho A (2004) Boosting for text classification with semantic features. In: International workshop on knowledge discovery on the web. Springer, Berlin, Heidelberg, pp 149–166
21. Tong S, Koller D (2001) Support vector machine active learning with applications to text classification. J Mach Learn Res 2:45–66
22. Kim SB, Han KS, Rim HC, Myaeng SH (2006) Some effective techniques for Naive Bayes text classification. IEEE Trans Knowl Data Eng 18(11):1457–1466

Blockchain for Management of Natural Resources Using Energy Trading as a Platform

S. Thejaswini and K. R. Ranjitha

Abstract Nowadays, the need and necessity of renewable energy are one of the most important techniques which play a major role in our day-to-day life. Renewable energy is one of the best natural resources. In the earth, energy is available in the form of light, breeze, rain, surfs, and thermal. Resources are naturally renewable and it can be renovated. The proposed work mainly concentrates on blockchain technology in rainwater management because the limitation of water is one of the major reasons which is mainly caused by the climatic changes in the weather, floods, and amplified pollution. Blockchain helps to make a perfect decision by the assortment of data of water quantity and water quality. It empowers clean water projects to tap investors around the world with unparalleled breadth and ease. By giving a strong fireproof cabinet, safeguards transparency, and distributed ledger to document relations among the festivities, this branch of knowledge might deeply indicate a change in the way of rainwater resources which can be achieved and traded easily and one example is the use of the smart water grid (SWG) which is a two-way mechanism. One of the best is network transmission between the sensors and the mechanism that is constant and distant used for observing the water supply mechanism. The other one is the smart grid that can display many alternative parameters such as coercion, an improvement in the quality, rate of flow, a temperature which is more efficient toward this technology. By making use of this technology, a model based on an energy trading system helps the growth of industries from small-scale industries to large-scale industries that provides outstanding practice for energy efficiency measures is proposed. It mainly helps in monitoring the truth value of water. By the use of smart grid and sensors, the data can be fetched within a fraction of seconds. Finally, blockchain collects the exact record of water transaction and this makes as an account which helps in the management of water resource.

S. Thejaswini · K. R. Ranjitha (✉)
Department of CSE, Siddaganga Institute of Technology, Tumkur, Karnataka, India
e-mail: ranjithakr.1si18scs08@sit.ac.in

S. Thejaswini
e-mail: thejaswinis@sit.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_36

Keywords Blockchain · Smart grid (SG) · Smart water grid (SWG) · Rain water · Machine learning (ML) · Intrusion detection system (IDS) · Deep learning (DL)

1 Introduction

India is the land of agriculture. Without water, there is no farming. Similarly, rainwater is a very important renewable energy asset because rainwater is always evergreen for the survival of plants and animals. It conveys new water to the earth's crust. If the scarcity of rainfall is less, then gradually the problem arises. Natural energy resources are solar power, wind power, small hydropower, biomass, etc., whereas renewable resources are natural resources which will regenerate naturally after consumption. But the time duration is high to refill or reload the resources. Hydropower generation is one of the examples of renewable energy which is able to generate few megawatt of power by converting the water potential into electrical energy with the help of turbines. Natural sources of water are categorized as rainwater, underground water, and surface water. Rainwater is another source which improves the water cycle above the surface and below the surface as groundwaters. Still large amount of water is present in a hidden state in different forms as glaciers and ice caps. This is how natural resources play an important role in the world. Rainwater is a mixture of electrolyte, abundant ions, sodium, and various minerals. Renewable energy resources are one of the main techniques which helps to convert from one form of energy to another form of the energy that is in the form of water cycle, water fades through solar energy and advances potential energy when the water hastens. This cycle of evaporation provides a mechanism for the conversion of solar energy into electrical energy. This rainwater can be stored as energy in many other forms. For example, in South Africa, the volume of water in storage is smaller in excellence due to the rise in inhabitants of a particular place and the area of the land where the water collects when it rains is surrounded by hills, truly generated by the modernization of urban, cutting of the plants and trees, damages of the rivers, the devastation of swamps, economic activity concerned with the processing of raw materials, obtaining coal or other minerals, from a mine, area of cultivation, usage of the work and unexpected poison of water pollution. This rain supports the blockchain technology in a greater way by providing a secure, transparent, and distributed ledger to record transactions between parties; blockchain-based technology [1] can primarily exchange the path of water utilities which can be succeeded and marketed easily. This smart grid technology helps in rainwater management in a greater way than by providing an electric grid which includes operations like smart meters, smart appliances, etc. It has a strategy called a smart grid policy which is systematized in Europe as Smart Grid European Technology Platform [1, 2]. Finally, this shows that rainwater is one of the renewable resources because rain is usually in a liquefied form which is molded as a result of condensation in the distinctive water vapors.

2 Renewable and Non-renewable Resources

Renewable energy sources are abundantly present in the nature such as solar, wind, hydro, thermal, and biomass energy and these energies are never depleting. The non-renewable resources such as coal, fossil fuel, natural gas, and petroleum do not successfully satisfy the needs of the people. But it is very efficient to use renewable energy resources than non-renewable. Nowadays, human beings are wasting water every day [2]. After all, they do not know the true value of water. And the rainfall has become reduced due to fewer forests, fewer trees, and deforestation which deliveries in the rainfall reduction. Water recycling is one of the best processes. This routine is almost done by the process of the flock and turns toward to form rain. The abrupt or immediate cause of a rainfall scarcity may be due to one or more issues including a non-appearance of available moisture in the atmosphere; large-scale collapsing which defeats convective activity; and the non-appearance or non-arrival of rainwater systems. Cutting down of trees will cut a significant amount of vapor from the atmosphere. This reduces the amount of rainfall received in the atmosphere.

3 Blockchain Technology in Rainwater Management

It is very hard to imagine our lives without water. Fine, 4 million people started their daily life's terrifying in Cape Town, South Africa, last year. Bureaucrats projected the town that there are only six periods of the time between the same dates in successive calendar months of investments is present. Through the shared process of doing something and some to a remarkable resultant rain in an area, they closely declined successively move away from water. The deadline for South Africa's water situation never ends. Weather conversion and deprived water system endures towards pay to water dearth, high-level pressures at activities, not to waste and quarter scale of the amount. In the next few decades, some of the major cities like Tokyo, Beijing, London, Mexico, Barcelona, and Istanbul are going to face day zero which represents zero level of water in the cities. Report states that in 2018 an average of 2.1 billion people is facing difficulty to access water for their regular needs.

The total volume of water is 1.386 billion km^3 loads of water on earth, in which 97% is of saline water and almost 2.5% remaining water is absorbed below the surface of the ground. The situation never goes roughly 0.4% to divide among one billion having 9 zeros human beings. It also concentrates on our whole commerce and cultivation sectors which provide a way for provisions toward adequate water to function. Weather reports change an industrialized organization at the last evolving world wannabe to devour like to walk or progress slowly into water scarceness. Expressing motion toward pillories is even worse, hominids are whispery celebrated their helplessness to share resources equally. Conflicts have been war nowadays for water during the legend. This sequence of actions continues until now. Blockchain technology can provide two ways of solutions which are as follows, water transparency and cooperation for smart water management.

3.1 Water Transparency

Blockchain is always safe, transparent [3], and scattered for concerning the record of events for communications across a formally constituted social gathering**.** If public blockchain is used for water and its excellence facts, then the evidence cannot be secreted or altered by the crooked behavior of politics, firms, or strong folks. The degree of excellence and the accent can be run through a conclusion in the period of growing water scarcity. Visualize ménages, industry, consumers, water supervisors, and policymakers use this evidence to conserve or use water. Activities of water supervisor are disposed to sleaze and devolved benefits, by giving the impression of being something perfect movement toward smart water system. Hong Kong has stated research toward water systems through WATERIG to identify rainwater through various points. These central parts of the wheel (hubs) are into contact with the water handling systems and they mainly focus on bids characteristics of vertical farming. Since a set of things is reorganized, humanities can come or bring to a resolution in which everything uses the blockchain to bulky account for their own colorless, transparent, odorless centers. WATERIG is chiefly contained with its colorless, odorless, and tasteless supply. This provides a ruinous study of Flint. About $28 million was given as a support by Flint's governor in the year 2016 to identify alternate to the natural calamity. Authorizing societies and slight firms to revenue toward rainwater encourages the market and networking system to a chance for employment. Rainwater under normal conditions gets wasted, but when societies provide openings to program their rainwater using the blockchain, this motivates the people toward recycling. The World Economic Forum recounted that rainwater might be cast-off to casual and influence factories or to create fertilizer. The municipal is contented and density is condensed on the system without special interventions. UN world record in 2017 reports that 80% of rainwater is wasted and does not treated for future use. As water becomes more insufficient for demand with weather change and growth of the world populace, societies will be further fortified to take benefit of their water gathering hubs.

3.2 Cooperation for Smart Water Management

Blockchain collaborating with the Internet of things (IoT) helps water systems cleverer, harmless, and more effectual. IoT centers for the operations of the water organization system, objects would look identical changed. The urban area water dispersal grid indicates the consequence of an imagined event of automotive and industrial radars will collect the records on the degree of the accent of water. Through online sensors communicate and investigate for seepages, channel torrents, and the state of making or being made impure by pollution or poisoning. But they swiftly send an alarm to indicate water executives. Based on the alerts, executives alter the pressure of water to avoid huge damage. In this view, Internet of things (IoT) provides better coverage for the betterment of every part of cities in a water interconnected system. The same knowledge is dropped in the backgrounds. Smart sensors interconnect one another to establish uncommon freshwater taken from the ground or surface. These observations help to see the water feasting visibly, choosing to select your performance and rescue currency. This alarms the mobile instantaneously for leaks even if the person is not at home and stops the water supply. In Internet of things (IoT) systems, all particular groups move from a lone view of safety intelligence. Entire systems conclusions are made here, producing it defenseless to equitation and maneuver. Just visualize, nowadays the water mechanization is erroneous in hands which gives rise to a threat to the national refuge. Blockchain accepts the undesirable failure in an Internet of things (IoT) scheme, by authorizing and securing the network. The combination of Internet of things and blockchain will be a fantastic job in water management.

4 Smart Grid in Rainwater Management

The smart grid is a shared term functional to an assembly of technologies, counting smart meters and meter reading apparatus, wide-area checking systems, lively line evaluation, electromagnetic autograph extent and investigation, time-of-use and physical stretch estimating tools, innovative shifts and cables, radio transportations technology, and numeral caring relays. A smartgrid [2] includes records. Each record has information that helps the clients to manage conveyance and detect incongruities and fiddling. Figure 1 represents the applications of the smart grid in various energy resources. The smart grid impression also embodies a trend in technology and the commerce prototypical for a convenience's connection with clients and extra sponsors. In the ancient model, a utility merely fashioned energy, transporting it over broadcast and spreading linkage to customers. In the new model, the helpfulness and its patrons become allies in dealing with the supply/demand connection. The smart grid provides the machinery to enable this connection. Using advanced metering infrastructure (AMI), smart meters have the power which regularly interconnects with the utility, electric efficacy can set charges which track costs that diverge by the time of day, swelling the price of power used during high demand periods, and decreasing it during truncated claim periods. Clients can retort to that material by altering their power weights, consequently, tumbling demand during peak usage periods, or buying utilizations that can retort to such information. For water utilities, there are foremost equivalents with power creation, transfer and buyer amenity, as well as significant variances. While there are major variances between water and electrical service, water conveniences involve many of the identical works and

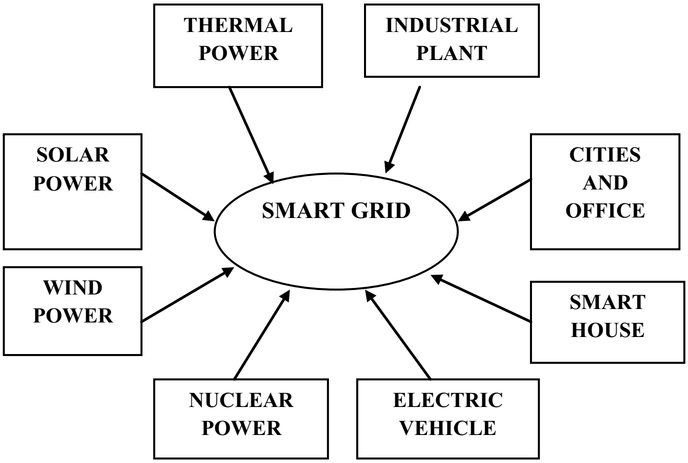


Fig. 1 Smart grid applications

topics that drive electric abilities, counting the necessity to grasp costs down while preserving system steadiness, powerful wealth necessities, the necessity for maintenance, conservation worries, and enrolment matters. Besides, water and wastewater benefits apply the foremost parts of their profitable resources on power. Mainly used for seriatim pump engines, which adds the measurement of responsive power. Vitality charges for water and wastewater can be tierce of a metropolis's energy bill. Portable water and drain water in the USA employ about US$4 billion a year on imparting the fluid energy, distribution, fetch, and sparkling water. Portable water and drain water conveniences justification for a predictable 75 billion KW/h of overall US power demand and tons are expected to grow by 20% over the next 15 years due to amplified populaces and more harsh guidelines. While water efficacies have more elasticity than electrical utilities in merchandise storage, many are theme to eventual anxieties that necessitate propelling when the cost of power is great. So, wastewater services stand to be jammed by the dawn of smart grid machinery.

4.1 User Feasting Outline Demonstrating a Leak

Water conveniences can use smart metering to display and device alternative day scattering vetoes. Water conveniences can launch periodic or regular period of usage or highest freight assessing, as well as tiered estimating for water resources. Such assessment is already being recognized in California and at other services around the world, expressly in thirstier zones. Smart metering offers the material to help clients retort to such worth gestures previous to getting the bill. In belief, a water value might send load resistor signs to irrigation managers and extra strategies, just as plug-in utilities would guide regulator signals to air conditioners and other tackle to shut off or dodge rotary on through certain periods. By using claim running techniques, a water efficacy could diminish the sum of water it needs to put into burden system storing (elevated tanks); this is tentatively equivalent to rechargeable efficacies dropping the sum of rotating fall-back that they obligate to retain operationally.

4.2 Need for Smart Grid in Water Utilities

Nowadays, a variety of operations and energy came into existence. So in rainwater management smart grid can be referred to as an apple because the one-half side is used to improve water infrastructure and the other half side is used for energy efficiency. Water conveniences will be a theme to more cultured assessing and claim side organization by their power providers. Smart grid retailers will adapt rechargeable smart grid machinery, particularly advanced metering infrastructure (AMI) and meter records organization systems, and arcade them to the water effectiveness arcade along with novel commercial models. Regulators will hearten water conveniences to assume power utility and commercial models. Their outcome will bear water convenience construction, dispersal, and client facility.

Smart grids are used to distribute the energy generated from renewable energy sources by connecting distributed energy sources. So this smart grid can be effectively used in renewable resources. One of the main reasons to include a smart system in the renewable resources is to gather information and communication technology (ICT) into the water network structure wherein which it increases the efficiencies of all elements in the water network. Smart grids support to manage the energy resources and reduces the power demand when any one of the energy sources gets dips by balancing the excess from other sources. Due to this energy management process, a smooth and efficient energy utilization could be achieved in an easy manner.

5 Major Problems of Water

In India, there are many water problems. Some of them are water conflict and water monopolies.

5.1 Water Conflict

Water conflict is a term relating to between republics, states, or assemblies over admittance to water incomes. A malicious process of water misrule devises fired radical disaster and controlled to humanity extremely feels that something cannot rely upon the central government's capability to resistor their water supply. This is an example of a clash between Karnataka and Tamil Nadu states in India for water conflict. Progressively thirstier region of the earth having the specified climatic conditions over the previous decades it has been igniting or cause to ignite again in the countries. Powerful combats generated gigantic objection as persons struggled for the scarcity of water which retains their folks active. The counterpole of the biosphere is the transformation of water by the Bolivian Government in 2000. This worsened into the "Water War of Cochabamba." After a broad conflict, the metropolis's water was re-nationalized and established the creation of lawful backing. However, on account of current water lack, some areas of the Bolivian countryside lost 90% of woodland during the drought in 25 years. For many societies and relatives, this is a matter of life and death. So the humble request is to save water and save lives which is most important for the biosphere.

5.2 Water Monopolies

Water is a natural resource that is controlled almost globally. Even a monopoly cannot charge any price it wishes. Once the patent expires, similar products will enter the market and the price of the product will fall. Monopolies are mostly measured to have numerous drawbacks (advanced value, fewer motivations to be effective).

By making use of rainwater as a renewable source of energy, the rainwater helps the energy storage device in a very efficient way. This rainwater comes in contact with the smart grid which includes records and bids to supply and investigate massive sizes of circulation and feeding information, offer it provides helpfulness work and clients to benefit them succeed transport and custom, and detect modifications and shifting. These routines a blockchain to provide transparency [3] and it will provide cooperations for the generation of smart meters.

5.3 Prediction

Machine learning models are widely used for prediction operations as a better prediction model provides better decisions based on analysis. In blockchain technology, machine learning models are used to predict the price of bitcoins. Valenkar et al. reported a prediction model using Bayesian regression analysis and random forest method to predict the bitcoin price through various parameters like size of the block, number of transactions, total number of bitcoins, and its trade volume. Normalization of trained dataset is achieved through box-cox, log and z-score. Recently, various price prediction models are evolved for predicting the values of Ripple, Dash, Ethereum, Litecoin, and Bitcoin cryptocurrencies.

6 Proposed Model

The energy trading system is proposed by using rainwater which is depicted in Fig. 2. Rainfall is one of the best factors of the water cycle and it is liable for leaving most of the renewed water on the earth. The situation offers opposite environments for several sorts of ecologies, as well as water for hydroelectric power plants and harvest irrigation. So this proposed model chooses the best renewable resources. When the air temperature falls to its dew point, water mist shrinks to form clouds. The air chills and the water mist shrinks, creating rain droplets. So this rainwater can be stored as energy. The concept of security of management of natural resources is proposed in which it includes air, water, soil, minerals, plants and animals are just used as a natural resource wherein which the humans use to produce the energy and make the things what the people use. The natural resources can be protected by cutting down on what you throw away, conserve water, plant a tree, choose sustainable, volunteer for cleans ups in your community and many more help to protect the natural resources. In this proposed model, a blockchain and smart grid for energy transactions using cryptocurrency is represented which is as in Fig. 2.

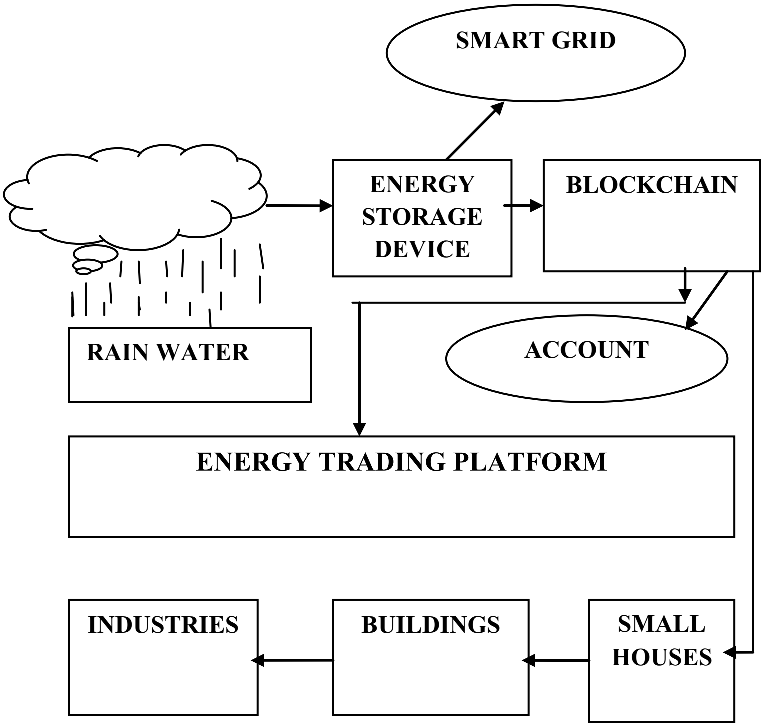


Fig. 2 Energy trading system using rainwater

It is more reliable for energy transactions. It uses deep learning (DL) and blockchain as a major resource [4, 5]. Byzantine fault tolerance is the feature of a distributed network to reach consensus even when some of the nodes in the network fail to respond or respond with incorrect information. In blockchain, byzantine fault tolerance can continue even if some of the nodes fail or act maliciously. Byzantine fault tolerance algorithm [6] is mainly used toward the yield of high structure speed. Blocks are created by hash functions and a dumpy signature. The proposed model consists of five ways infrastructures: buyer, seller, blockchain, intrusion detection system (IDS) [7], and machine learning [1].

The buyers trade the vitality with the seller. The buyer displays sufficient cryptocurrency which is required to purchase the sufficient asset. The buyer can neither be a person nor a building, so that the buyer could be situated in a dedicated network connecting devices in the home such as display, load control devices and finally smart applications or it can be a collection of smaller local area networks. Seller always gives rise to energy from renewable resources which include tidal energy, wind energy, biomass, etc. The seller can be a fellow citizen. Blockchain in rainwater management has more advantages for water scarcity. In this proposed system, the blockchain clearly collects the data of water quantity, and blocks are generated through its short signatures and hash functions. Based on recurrent neural network, this intrusion detection system works and identifies the fraud transactions and attacks. Further quality and efficiency can extend to make an account. This collection of data can be used to evaluate data analysis [8] which provides a platform for energy trading platforms. Energy storage device uses a smart grid because grid management is always cost saving which as a significance toward water management. Small houses, buildings, and industries are the major examples because it collects the data in an efficient way, and this helps the blockchain to list the records in a very efficient way. This unit is a distributed digital ledger [9], by which it is surrounded by all vigor relations in the smart grid system. Seller produces a vigor from the renewable energy resources which is used for itself and in the second step seller left the energy resources. Finally, the seller distributes per element fees which is a measure of power liveliness on the blockchain network. In peer-to-peer (P2P) [10], the energy trading system, the desirable energy buyer resolve aspect on the top of the printed worth. Frequently all the energy buyer drafts his justification which takes sufficient cryptocurrency stability for energy trading. Formerly, if the vigor claim of energy buyers and per-unit price coordinated their condition of energy at that instant, it pushes an acquisition appeal to the energy seller continuing in time the blockchain network. Once the successful authentication is performed, it provides option to buy energy resource through cryptocurrency. In this peer-to-peer (P2P) [10], energy trading system obtains the seller and buyer details in order to avoid hurting transactions. A spiteful block or lump can impress the peer-to-peer (P2P) system and contest for energy trading.

An energy trading platform has the ability to transform the energy from one form to another form. The proposed approach requests the buyer to consider the price per unit of energy and checks buyer account has sufficient credits for trade. Based on the demand of user and amount, a purchase request is forwarded to the service provider through blockchain network. Upon validation of buyer, seller provides resource to the user and accepts cryptocurrency. This P2P trade requires data analysis in order to identify the frequent buyer and seller, so that malicious transactions could be identified in the network as it will produce huge impact to the P2P system and energy trade. Blockchain is a decentralized security system that prone to attacks [11]. The important factor to be considered in blockchain is its consensus protocol, so that the blocks are controlled by few farms. Public blockchain has these features, whereas private blockchains are not affected by the attacks. Since each node is identified in this network through suitable consensus protocol, blockchain provides better control and efficiency in energy management. Also, real-time update and security features are provided through distributed ledger.

Intrusion detection system (IDS) [7] can figure out the untrustworthy connections and connection occurrences on the energy trading platform. The intrusion detection system mainly helps to alert the teams when leaks have defected. The intrusion detection system can also be set to detect water leaks, fire, an open window, or any other anomaly which may put property or personnel at risk. In intrusion detection network, intrusion detection systems can be observed in this proposed system because they easily identify the anomalies with the aim of catching the hackers before they do real damage to a network.

Machine learning (ML) [1] can be implemented to blockchain technology to make the petitions smarter. By making use of machine learning, distributed ledger can be maintained with high security. This machine learning supports to obtain suitable paths for data sharing. Various sources of data are collected and processed in the smart applications and blockchain will be a part of this smart application. Machine learning can be used to analyze the collected information and predicts the values and the collected information are stored in blockchain network. This avoids the issues such as redundant data, missing values, noise, and errors. Most of the data-related issues could be eliminated through blockchain. Integrating machine learning with blockchain classifies the essential data instead of analyzing the entire information in the dataset. Based on this, various approaches and applications could be developed in future for fraud detection.

7 Conclusion

Advancement in blockchain technology made a tremendous path in the technical environment. In this proposed work, detailed information on blockchain technology in the smart grid by using renewable resources that are rainwater. The architecture mainly represents energy trading system for data analysis using deep learning and blockchain technology. The energy trading system plays a major role which helps in the storage of energy using renewable resources. In this proposed work, rainwater is the best natural resource, this energy is stored in an energy storage device, and this energy storage device comes in contact with the smart grid because of smart meters and sensors which provide a dynamic way of monitoring the system. A distributed ledger is one of the main roots of the smart applications such as smart health care and smart grid. Making use of this smart grid (SG) and blockchain can evaluate the data analysis using an energy trading platform. This major activity helps the industries, buildings, and small houses through account so that it is greatly benefited in an easy manner. In future, this account can be replaced by the digital applications and since the paperwork concentrates mainly on the proposed model further this can be extended to implement it efficiently.

References

1. Tanwari S, Bhatiai Q, Patel P, Kumari A, Singh PK, Hong W-C (2016) Machine learning adoption in blockchain-based smart applications: the challenges, and a way forward
2. Bracciali A, Chatzigiannakis I, Vitaletti A (2019) Citizens vote to act: smart contracts for the management of water resources in smart cities
3. Abe R, Watanabe H, Ohashi S (2018) Storage protocol for securing blockchain transparency
4. Sun L, Chen X, Yang Z, Ke D, Meng, Qiu J, Cao Y (2019)Energy storage strategy in a non-agent energy trading platform: energy bank system
5. Wu H-T, Lu C-Y (2019) A deep learning application system based on blockchain technology for clicks-and-mortar businesses
6. Matsumoto Y, Kobayashi H (2010) A speculative byzantine algorithm for P2P system
7. Borkar A, Donode A,Kumari A (2017) A survey on intrusion detection system (IDS) and internal intrusion detection and protection system (IIDPS)
8. Nugent M, Lennon RG (2019) Blockchain for decentralized data analysis
9. Zhang K, Jacobsen H-A (2018) Towards depenadable, scalable and pervasive distributed ledgers with blockchains
10. Dorriab A, Hillab A, Kanherea S, Jurdakb R, Luoc F, Yang Z, Donga (2019) Peer-to-peer energy trade: a distributed private energy trading platform
11. Sivaganesan D (2019) Blockchain enabled internet of things. J Inf Technol Dig World

Machine Learning Model for Anomaly Detection in Big Data for Health Care Applications

M. G. Sharavana Kumar and V. R. Sarma Dhulipala

Abstract Recently, enormous amounts of data are increased by the essentials of data security and investigation for big data. Anomaly detection system monitors the data to analyze and detect any intrusion/misbehave in the network or devices. With the traditional detection techniques, it is highly complex to perform analysis and detection process. Techniques for big data are widely incorporated for anomaly detection system in achieving efficient data analysis. Hence, it is attempted to introduce fuzzy logic-based anomaly detection. The efficiency of the result has been obtained with greater accuracy with high performance with the added advantage of reduced training time.

Keywords Big data · Anomaly detection system · Machine learning · Fuzzy logic

1 Introduction

In the last decade, internet applications and communication have been tremendously developed in the medical field of information and communication technology. An application that falls in that category could generate huge generation, various difficult multifaceted structured data usually known as big data [1]. Since this era is full of automatic data collection, systematic measurements could not even know about the relevant information [2]. An example of this kind of application is E-Commerce; every transaction in E-Commerce includes buying, selling, and investing [3]. Hence, it generates health related data with complex structure to process and store which leads to the situation that traditional data storage techniques become fatal. This leads to handle the data and analytics it enables big data into the platform [4]. Nowadays, wearable sensors and systems along with people use the web with an exponential generation of the massive size of data, size of the data might be measured in exabyte

---

M. G. Sharavana Kumar · V. R. Sarma Dhulipala (✉)
University College of Engineering BIT—Campus, Anna University, Tiruchirappalli, India
e-mail: dvrsarma@aubit.edu.in

M. G. Sharavana Kumar
e-mail: mgsharavanakumar02@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_37

(EB) and petabytes (PB). It is expected that in 2025, there might be some yottabyte [5]. Roger Magoulas a researcher who coined the term big data for the first time to describe the propensity.

The expanded amount of health data provides a general overview of large databases through big data systems and computational processes [6]. Large data are semi-structured and unstructured data relative to standard data sets which tend to be more real-time analyzed [7]. Big data allows to define emerging principles and permits the fully grasp which is the secret ideals in greater detail, although it also raises new obstacles. Unanticipated massive cloud and internet computing (IoT) growth advanced the growth of data collected from medical devices [8]. Cloud infrastructure is now well established to provide for the collection of data and business connectivity for massive computer holdings [9]. In IoT, wearable medical sensors require the data to be obtained and data sent to the cloud to be analyzed further. These data usually surpass existing processing power and contributes to issues with the collection or reservation for the unique hardware and software infrastructures of large heterogeneous datasets.

Big data is difficult to store due to the rapidness and problematic to achieve and examine dynamically using the traditional storage techniques both software and hardware [10]. Big data has its characteristics of high volume and velocity, data variety, which needs emerging techniques to handle it. Anomaly detection system (ADS) monitoring is both hardware and software monitoring that analyze and detect data by any form of attack in system or network. Existing old-style anomaly finding techniques is complex and not up to the level of efficient, when it happens to be big data [11]. Time constraint is the major thing to deal with, since existing system is prone to harm due to time delay, there is a need to introduce the efficient techniques to store and analyze the data in ADS. Thus, the anomaly detection calculation time and alerting period can be in actual/on time.

ADS has three methods for detecting attacks namely, signature, identification of anomalies, and combination detection. Signature-based uncovering is one of the effective methods to detect in such a way that preloaded in the ADS database, thus it will be accurate enough to identify the attempt of an anomaly in a known type of attack [12]. The disadvantage of this detection mechanism only pre-defined attacks can be detected it cannot detect the new type of attacks, since its signature not in the anomaly database to increase the efficiency of the detection the database needs to be updated frequently/periodically [13]. Even the updated database effectiveness is not up to the level, since it has a high false-positive rate.

The hybrid-detection mechanism is the combination of two or more detection mechanism of anomaly detection enables the hybrid-detection technique. To overcome the pitfalls in the mono detection mechanism, it is preferred to use a hybrid-detection mechanism [14]. To handle the big data along with machine learning is one of the effective methods ADS is for anomaly detection to reduce the false-positive rates and to increase the effectiveness of the detection accuracy.

2 Methods and Background Studies

Machine learning (ML) is an emerging field in the research area that primarily focused on the theory and performance of the algorithms [15]. Those algorithms are highly interdisciplinary filed such as artificial intelligence, optimization, data, knowledge, and commerce. Almost any technological area, which has already had a significant influence on study and development, has already covered the way the data are treated in a broad variety of applications, Machine learning, and artificial intelligence. This was used for different issues, including engines. Machine learning is generally divided into three categories: supervised learning, uncontrolled learning, and enhanced learning as shown in Table 1.

The huge amount of data in real-time has hindered network analysts' success in scaling up the plentiful data rate. It is necessary to generate actual and near-in realistic, and not from the periodic log data available on the network protection review [16]. The current anomaly detection technologies, however, are unable to store, identify, and report on the real-time or near real-time data [17]. Additionally, more and more different attack styles are discovered every day in networks, but the existing detection systems have struggled to identify these assaults due to the immense number, size, range for study.

In the era of big data, the network security issues have been many and still not covered, particularly in the areas of modeling, data analysis, and anomaly detection [18]. This includes the fields in which the data process, banking, education, and transport are still not addressed. However, conventional surveillance systems that exist do not manage large data and are thus unable to constantly track the network architecture and recognize irregularities and vulnerabilities are necessary for rapid analysis of the data generated to recognize possible risks to the network [19]. Thus, a fusion of deep learning algorithms combined with large data technology is necessary to overcome the holes in the bottlenecks of current techniques and to effectively

Table 1 Methods and characteristic of big data

Types	Methods	Characteristic norm	Name of the algorithm	Types of data	Approach
Supervised learning	Classification & Regression	Computational classifiers Statistical classifier	SVM Naïve Bays	Labeled data	Maps the labeled inputs to the known outputs
Unsupervised learning	Cluster formation & prediction	Parametric classifier	K-means neural networks	Unlabeled data	Understands patterns & discovers the output
Reinforcement learning	Decision making	Modeling & non-modeling	Adaptive learning based	No Predefined data	Trail & error method

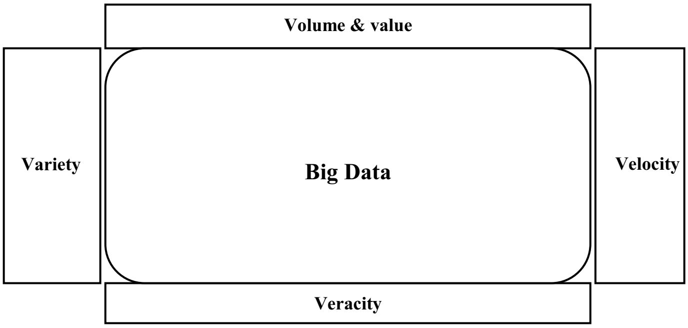


Fig. 1 Key issues in big data

process large real-time data to identify anomalies. Machine learning knowledge helps to detect using algorithms by analyzing the data collected [20]. With the help of back propagation algorithms to train the RNN, the work projected in such a way to improve the system reliability by using the SVM algorithm, thus the robust prediction can be ensured for different applications [21]. Besides, various big data technologies can contribute to the real-time and almost real-time processing and streaming of enormous amounts of network data. There are many issues faced by machine learning techniques such as large-scale data, various types of data, high-speed streaming data (including multimedia data), uncertain and incomplete data, and low-value density data as shown in Fig. 1.

The various terms referred to in Fig. 1 are namely, volume—large scale, variety—heterogeneous, non-linear, velocity—real-time and high-speed stream, veracity—uncertain, value—low value, and diverse.

3 Challenges of Anomaly Detection System (ADS)

There are challenges for ADS devices that are deployed on a large network is in the ADS components communication in and around of the sub-networks, in some case finished firewalls and entries also. Different amounts of the system and different network devices their data formats which are usually in unique and different protocols with different data formats. Another major challenge for ADS in a huge network is effective traffic monitoring. Network anomaly detection system (ADS) uses many different components that are speckled through a system, if those components are not placed in a specified way, many attacks may intrude in network anomaly detection system (NDS) as shown in Fig. 2.

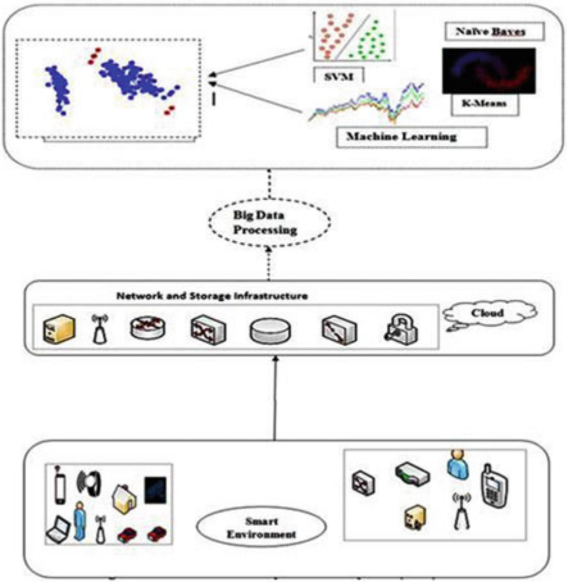


Fig. 2 Flow chart of anomaly detection system (ADS)

3.1 *Fuzzy-Based Detection Factors*

The concentrated logic detection block makes choices concerning healthcare and the presence of anomaly and therefore proposes a suitable fuzzy level implementation. It is carried out based on human logic, which imitates the same strategy of reasoning. The intensity of the system and progression of the system are based not on personal health, but also on the complete detection of an anomaly. In the fuzzy systematic evaluation logic, there are overlapping fluctuating changes in the interval which mimic the way people perceive modification along with reasoning. Fuzzy logic describes the way of making use of fuzzy if/then or or/and rule of imprecise dependencies and commands [22].

The linguistic variable, membership function, and the speech are the concepts that play a vital role in the application of the systems. A fuzzy set is a set of real numbers with part of the set. The layers of the group are calculated by a value of membership function, absolute non-members in 0 functions in the system, and half of the community by a value of 0 to 1. The MF is a curve that sets the form of benefit in a matrix for each point inside the input space. The member's role is for each input to find the standard range, with MF specifying a set.

Consider the scheme under which a vector in a discourse of anomaly and is an actual number in which denotes the fuzzy collection described. The member function associated with these conditions is a function that maps into [0, 1] and provides the function rank of anomaly detection. The MF used in the fuzzy logic system is triangular, provided that it is within [0, 1]. It then defines the function of the deltoid anomaly function in which the parameters are on the input the fuzzy logic works to achieve the crisp output α.

A multiple input and one output structure are built for the study of health performance. The model is designed with several entrances and every if-then rule is defined by fuzzy input sets and output set of parameters.
The linguistic variables are defined by the value of the output. Accordance with the data of the fuzzy system, the de-fuzzing framework is necessary, to get the crisp out of the method, since the output of each assumed fuzzy group is independently measured by a characteristic value of classification.

3.2 Dataset

KDD99 dataset is widely used for anomaly detection system performance evaluation. The dataset consists of nearly 5 million records including normal and abnormal events that are already collected to differentiate the anomalies synthetic anomalies has been injected and tested. İt is taken 5000 samples of data to check and validate among the huge number of records.

3.3 Model Classification SVM

One of the supervised learning methods in support vector machine (SVM), it examines data for organization and reversion. Usually, SVM categorizes data into divergent lessons. In the binary case, this method does classification in such a way that given data into a couple of classes with the help of a linear hyperplane. Consider a scenario if suppose a vector x exists, scalar q and r as shown in Eqs 1 and 2.

$$x^{\mathrm{T}} x + q \geq 1 \tag{1}$$

$$x^{\mathrm{T}} x + r \leq 1 \tag{2}$$

where $x-$ the weight vector is q and $r-$ is the bias value

SVM facilitates to reduce the error rate in classification, maximizing the margin. It achieves better execution when maximizing the margin involving the vectors of the two classes, which could result in the classifier in the rate of maximum margin. Equation (3) is used to find the best possible separating hyper lance

$$w = x, y = b, w = x, x = a, c = b, a = q, b = r \tag{3}$$

$$a = q, b = r, \tag{4}$$

$$\min \frac{1}{2} ||x||^2 \tag{5}$$

Subject to

$$b_i(xa_i + c) \geq 1; \quad \forall (a_i, b_i) \in D \tag{6}$$

In order to reduce the possessions of outliers and errors of misclassification, soft margin has been used in SVM. It enables the show in Eq (7) for the non-negative slack variable which is used to understand the trade-off between the margin and misclassification error, which usually end-user entity defined.

$$\min \frac{1}{2} ||x||^2 + C \sum_{i=1}^{O} \mathcal{E}_i \tag{7}$$

Subject to

$$y_i(wx_i + b) \geq 1 - \mathcal{E}_i \quad \mathcal{E}_i \geq 0, \quad 1 = 1 \ldots N \tag{8}$$

where $\mathcal{E}_i$ signify slack variable, $C-$ indicate a consequence parameter usually reins the trade-off between the cost of classification margin and errors of misclassification.

3.4 Genetic Operators

Efficient optimization takes place when genetic operators are taken into account, which includes basic genetic operators. In order to optimize the parameters, an efficient strategy needs to be incorporated. The genetic algorithm has the adaptability to fine-tune in the dynamic fashion of altering that operator with respect to the evolutionary situation in the population, in order to retain the variety of the populace and to connect early and enhance the exploration.

3.5 Selection Operator

Sort comparative technique has been used to verify the value that is in the fitness of every individual. The fitness value is proportional when each probability of selection and the corresponding value of fitness is proportional. If the group size is n, if the adapt degrees of the individual is f_i, probability pi which can be chosen with the help of the following formula.

$$Pi = \frac{f_i}{\sum_{i=1}^{N} f_i} \tag{9}$$

Cross operator intended of a grouping of linear vectors combined to numerical the value of the character. In case two individuals $T_b T_c$ cross, then the result of spring by an entity in the following formula.

$$T_b^{'} = \lambda T_b + (1 - \lambda) T_c \tag{10}$$

$$T_C^{'} = \lambda T_c + (1 - \lambda) T_b \tag{11}$$

where λ is a random number between (0, 1)

4 Role of Mutation

Using the probability Pm mutation operator randomly selects a value then it adds the value of entering point, it applies to all offspring chromosomes. Whenever GA into the technique, there are two considerations to take care of, primarily to prevent the premature phenomenon of population diversity in the initial mutation operation that can be large. Secondary consideration, a secondary consideration is the near-optimal solution neighborhood, how to make reduced the variations operator. To ensure the optimal solution, random search capability and acceleration of convergence are to be ensured. Hence, the adaptive mutation probability in the following formula solve the above-said problems.

$$Pm = \frac{\exp(-1.5Xt/2)}{\text{pop\_ size}X\sqrt{L}} \tag{12}$$

The ADS neural network anomaly identification parameters are the number of neurons in cached layers, learning duration, epochal amount, and the momentum of learning as shown in Fig. 3. Depending on the issue, the number of secret layers may be increased. An ADS with three secret layers is therefore sufficient to map every continuous feature to satisfy the necessary complexity by inserting a certain number of neurons. This is why the amount of secret layers used in our experiments is 3. Training pace and momentum are two main metrics for the effective training of the ADS network.

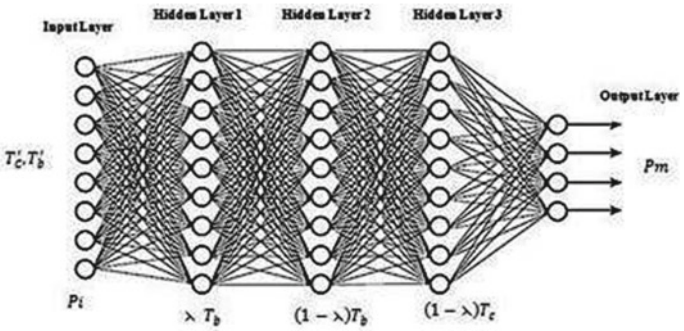


Fig. 3 ADS neural network anomaly and parameter

5 Result and Discussion

Changing parameters for neural networks in hidden layers is the distribution of neurons: [12,… 35] The first step is to be considered as a step with phase 3 parameter for the pace of learning: [0.03, 0.06, 0.08, 0.1, 0.5], for the number of epochs: Only one parameter change principle with fixation of all other parameters is created to search neural structure parameters.

Big data is used for the neural network detection of anomalies. Results are achieved by changing network parameters similar to the results of ADS neural networks with help of MATLAB software environment. The parameters used are the number of neurons, rate of education, number of epochs, and momentum in hidden layers. Big data's initial parameters are identical to ADS' initial neural network (the secret layer of neurons is 15; hidden layers 3 remain, and so on). Some of the improvements are still based on the same idea. The epoch variations [60… 300] and the amount of momentum [1.0… 0.5].

The network parameter modifications were collected for comparison of data-driven mapping neural networks (SVM, EB, and ADS). The problem is the global optimization function in which each goal is described as a shift in the given neural network parameter, the optimum forehead is present in Table 2. ForSVM, the distribution coefficient is just one target in building a neural network structural search space. All the scales (Accuracy, Performance, and Efficiency) have been measured as shown in Fig. 4. Instead of the neural network ideal parameters, the more complicated approach called the neural network architectures are considered based on the investigation.

The ADS provides great outcomes for all output indicators for both chosen pieces. Namely, for the chosen best parameters of EB modeling, the total precision scales (Accuracy, Performance, and Efficiency) by 98.82% as shown in Fig. 5. The ADS produces marginally better results (92–99%) compared to the neural SVM network (84–96%) with test information.

In comparison with the section (92–99.32%), neural network types (SVM and EB) give better rating results for the steam drums section (88–99%) as shown in

Table 2 Comparison of accuracy, performance, and efficiency

S.NO	SVM			EB			**ADS**		
	Accuracy	Performance	Efficiency	Accuracy	Performance	Efficiency	**Accuracy**	**Performance**	**Efficiency**
1	87.25	90.12	82.36	91.23	92.36	94.52	95.62	96.12	98.25
2	88.62	91.23	84.62	93.25	93.75	95.12	96.12	97.62	97.36
3	89.32	94.62	88.62	93.75	91.62	94.62	97.65	98.36	98.61
4	91.62	96.52	92.65	94.62	94.65	96.72	98.25	98.98	99.65
5	94.62	95.62	94.67	95.62	96.75	97.12	99.42	99.12	99.87

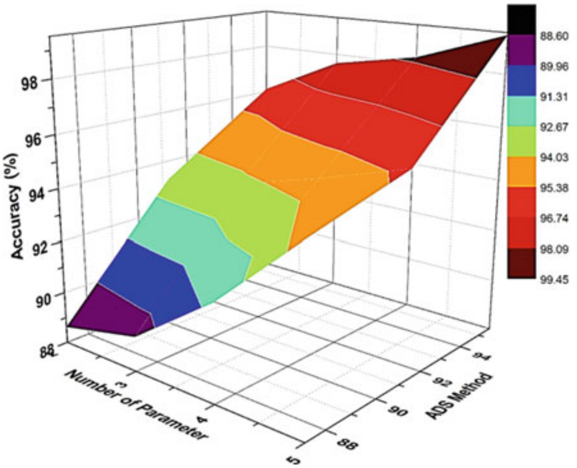


Fig. 4 Comparison of accuracy in ADS method

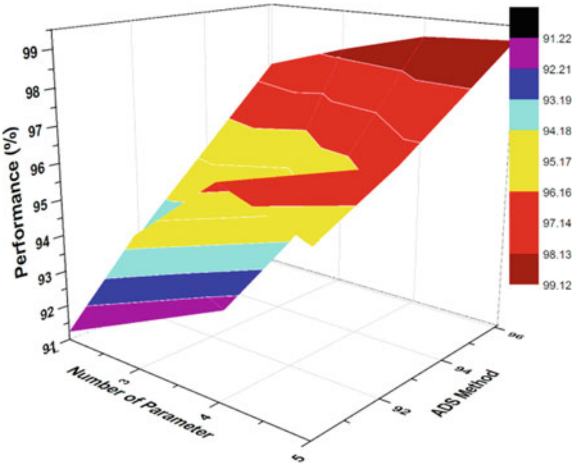


Fig. 5 Comparison of performance in ADS method

Fig. 6. For the steam drum portion, the accuracy and performance values are higher than the segment in specific parameters. It is anticipated mostly as there is a greater gap between data that is anomalous and data that is usual for steam drums. That is predicted.

The identification of abnormalities is a significant issue, which has been studied in numerous fields of study and implementation. Thanks to continual use, manufacturing systems become affected and this should be identified as early as possible to prevent losses. This is also important to provide a good device anomaly detector due to how they perform analysis, and the detection process method is one of the most complicated fluid structures that will still work appropriately at a minimal expense. Many fatalities have been attributed to heart attacks for the last ten years. Many scientists have developed many techniques of data-mining to diagnose these cardiac

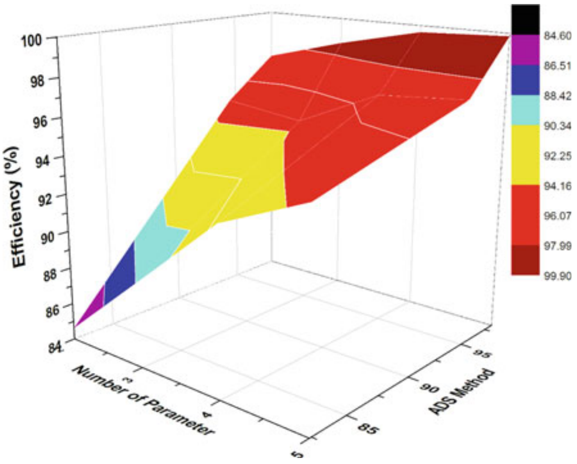


Fig. 6 Comparison of efficiency in ADS method

disorders for medical practitioners. When coping with cardiovascular disorders, the most popular methods for machine learning are ADS. Within this article, the procedure is used to treat cardiac disease, which incorporates k-means clustering and ADS. The findings of the trial indicate better specificity in the treatment of heart failure. Our process is 99.82 percent accurate; it is much stronger than many older methods.

6 Conclusion

With the conventional detection procedures, it is exceptionally complex to execute the investigation and detection process. Procedures for big data are broadly consolidated for a neural network framework in the discovery of anomalies. Machine learning (ML) is a growing field in the research of regions that fundamentally centered around the theory and performance of the algorithms. KDD99 dataset is generally utilized for the anomaly detection system execution assessment. One of the supervised learning methods in support vector machine (SVM), inspects information for identification and inversion. This work analyzes different models to investigate anomaly detection choices in selected parts of health care. Experimental results demonstrate that early anomaly detection in neural networks is highly efficient, with 99.52 percent overall precision (Accuracy, Performance, and Efficiency) modeling parameters.

References

1. Habeeb RA, Nasaruddin F, Gani A, Hashem IA, Ahmed E, Imran M (2019) Real-time big data processing for anomaly detection: a survey. Int J Inf Manage 1(45):289–307 (Apr)

2. Abdelghafar S, Darwish A, Hassanien AE (2020) Intelligent health monitoring systems for space missions based on data mining techniques. In: Machine learning and data mining in aerospace technology 2020. Springer, Cham, pp 65–78
3. Kumar MS, Dhulipala VS, Baskar S (2020 Jun) Fuzzy unordered rule induction algorithm based classification for reliable communication using wearable computing devices in healthcare. J Ambient Intell Humaniz Comput 24:1–2 (Jun)
4. Yuan G, Zhang C, Hu S, Guo J, Wang X (2019) Big data based bridge anomaly detection and situational awareness. In: 2019 Chinese automation congress (CAC) 2019 Nov 22. IEEE, pp 3864–3868
5. Garg S, Kaur K, Kumar N, Rodrigues JJ (2019 Jan 16) Hybrid deep-learning-based anomaly detection scheme for suspicious flow detection in SDN: a social multimedia perspective. IEEE Trans Multimedia 21(3):566–78 (Jan 16)
6. García NM (2019) Multi-agent system for anomaly detection in Industry 4.0 using Machine Learning techniques. ADCAIJ: Adv Distrib Comput Artif Intell J 8(4):33–40
7. Haripriya AP, Kulothungan K (2019 Dec 1) Secure-MQTT: an efficient fuzzy logic-based approach to detect DoS attack in MQTT protocol for internet of things. EURASIP J Wirel Commun Netw 2019(1):90 (Dec 1)
8. Abdelaziz A, Salama AS, Riad AM, Mahmoud AN (2019) A machine learning model for predicting of chronic kidney disease based internet of things and cloud computing in smart cities. In: Security in smart cities: models, applications, and challenges 2019. Springer, Cham, pp 93–114
9. Parvin P, Chessa S, Kaptein M, Paternò F (2019 Jan 1) Personalized real-time anomaly detection and health feedback for older adults. J Ambient Intell Smart Environ 11(5):453–69 (Jan 1)
10. de Sousa LD, Giommi L, Tisbeni SR, Viola F, Martelli B, Bonacorsi D (2019) Big data analysis for predictive maintenance at the INFN-CNAF data center using machine learning approaches. In: Conf of Open Innovations Association (FRUCT), Helsinki 2019, pp 448–451
11. Ahmed H, Younis EM, Hendawi A, Ali AA (2019) Heart disease identification from patients' social posts, machine learning solution on Spark. Future Gener Comput Syst (Oct 5)
12. Ramasamy B, Hameed AZ (2019 Jul 1) Classification of healthcare data using hybridised fuzzy and convolutional neural network. Healthc Technol Lett 6(3):59–63 (Jul 1)
13. Goyal KK, Paray AH (2019) A survey of different approaches of machine learning in healthcare management system. Int J Adv Netw Appl 11(03):4270–6
14. Larriva-Novo X, Vega-Barbas M, Villagrá VA, Rivera D, Álvarez-Campana M, Berrocal J (2020 Jan) Efficient distributed preprocessing model for machine learning-based anomaly detection over large-scale cybersecurity datasets. Appl Sci 10(10):3430 (Jan)
15. Yousefi S, Derakhshan F, Karimipour H (2020) Applications of big data analytics and machine learning in the internet of things. In: Handbook of big data privacy 2020. Springer, Cham, pp 77–108
16. Abd Ghani MK, Mohammed MA, Arunkumar N, Mostafa SA, Ibrahim DA, Abdullah MK, Jaber MM, Abdulhay E, Ramirez-Gonzalez G, Burhanuddin MA (2020) Decision-level fusion scheme for nasopharyngeal carcinoma identification using machine learning techniques. Neural Comput Appl 32(3):625–38 (Feb 1)
17. Razzak MI, Imran M, Xu G (2019 Mar) Big data analytics for preventive medicine. Neural Comput Appl 16:1–35 (Mar)
18. Selvaraj A, Patan R, Gandomi AH, Deverajan GG, Pushparaj M (2019 Nov) Optimal virtual machine selection for anomaly detection using a swarm intelligence approach. Appl Soft Comput 1(84):105686 (Nov)
19. Qi B, Zhang P, Rong Z, Li C (2020 Oct) Differentiated warning rule of power transformer health status based on big data mining. Int J Electr Power Energy Syst 1(121):106150 (Oct)
20. Bayrak EA, Kirci P (2019) Intelligent big data analytics in health. In: Early detection of neurological disorders using machine learning systems 2019. IGI Global, pp 252–291

21. Raj JS, Ananthi JV (2019) Recurrent neural networks and nonlinear prediction in support vector machines. J Soft Comput Paradigm (JSCP) 1(01):33–40
22. Kumar MS, SarmaDhulipala VR (2020) Fuzzy allocation model for health care data management on IoT assisted wearable sensor platform. Measurement 108249

Pilot Decontamination Algorithm with Iterative Weighted Graph Coloring Scheme for Multi-cell MIMO System in 5G Applications

V. Baranidharan, S. Karthikeyan, R. Hariharan, T. Mugunthan, and S. Vhivek

Abstract A multi-cell massive multi-input multi-output (MIMO) technology is the most important and promising technology in 5G wireless systems. The base stations in this 5G system are always provided with an abundant number of antennas. These base station antennas are simultaneously served more data to the set of users. Due to this more number of high-speed users, there will be high data traffic, the antennas suffer different types of pilot contamination in the various adjacent cells. A modified EIWG-WGC-based pilot contamination scheme is proposed to mitigate and avoid the unnecessary pilot contamination from the near adjacent cells. In the first phase, in order to find the potential PC, relationships among the users will be calculated by the construction of a modified edge-weighted graph coloring algorithm (EWIG). This graph coloring helps to find whether the two users are connected in the different cells by the weighted edge. This also indicates potential PC available and its strength for reuse of the same pilot. After the construction EWIG, the WGC-based pilot contamination method is widely used to segregate the pilot by assigning colors and finding the vertex of each user. This segregation helps to mitigate the PC by assigning weights based on different pilots. This proposed scheme is simulated by using MATLAB software. The simulated results show that this proposed scheme (EWIG-based WGC-PC) outperforms the existing system.

V. Baranidharan · S. Karthikeyan · R. Hariharan (✉) · T. Mugunthan · S. Vhivek
Department of Electronics and Communication Engineering, Bannari Amman Institute of Technology, Sathy, India
e-mail: hariharanr.ec16@bitsathy.ac.in

V. Baranidharan
e-mail: baranidhar@hotmail.com

S. Karthikeyan
e-mail: karthikeyan@bitsathy.ac.in

T. Mugunthan
e-mail: mugundhan.ec16@bitsathy.ac.in

S. Vhivek
e-mail: vhivek.ec16@bitsathy.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_38

Keywords Edge-weighted graph coloring algorithm (EWIG) · Multi-input multi-output (MIMO) · Weighted graph coloring-based pilot decontamination (WGC-PD)

1 Introduction

In the 5G wireless communication system, Multi-input multi-output (MIMO) systems are widely analyzed to meet the rapid and exponential increase in a large number of users and their high data traffic. In this 5G mobile wireless communication, the base stations (BS) are provided with an abundant number of transmitting and receiving antenna for supporting the more number of the set of users and support large data traffic [1]. If the number of antennas used is always fixed in the 5G massive MIMO wireless systems in which it gives BS as the total number of the antenna will goes reduce the inter-cell interference between the set of users and uncorrelated noise among the user's channels [2].

By the exponential increase in the number of antennas, it will lead to a massive MIMO system. This massive MIMO system is giving more benefits in the context of received signal detection scheme from the different antennas and estimation of the desired signal [3]. In this context, the contamination of pilots will be caused for the effective reuse of various pilots in the adjacent cells because of the limitation of less number of pilot's contaminations in a system. This pilot resource-effective utilization does not diminish the size of the BS antennas and values. So, the pilot contamination is to be recognized properly and it is being bottlenecked of multi-cell massive MIMO system [4].

There are numerous effects that have been taken to resolve these PC problems in these massive MIMO systems [5, 6]. These are different new methods that are evolved to solve PC issues. They are, time sifting pilot scheme: This scheme is more suitable for asynchronous transmission among the adjacent cells. Smart pilot assigning scheme: In this method, the pilots are assigned sequentially for each and every cell. PC precoding scheme uses multi-cell joined signal processing methods in a greater way. AOA method will assign the pilots based on a geographically separated by non-overlapping AOA values.

In all these existing systems, the severity value PC may carry for all the users and all this solution is also to try to mitigate and avoid the pilots for all the users, while the other will enjoy to negligible PC. In this proposed work, the modified EWIG–WGC-based reweighted pilot contamination scheme is designed and proposed to avoid and mitigate the contaminations of the pilots for this various multi-cell massive MIMO system is formulated the optimization problem in order to maximize and convergences of the uplink achievable rate. In the second phase, the modified EWIG–WGC construction is formed from the classical pilot assigning and contamination algorithms of Multi Cell massive MIMO systems. This algorithm greedily comprises and assigns a different pilot connect to multi-users for large iteratively re-weight pilots over the users.

2 Literature Survey

Many researchers have taken many efforts to resolve challenging pilot contamination (PC) problems. Some of the recent related works are explained in this section.

Wei et al. [7] have suggested the PARAFAC-based channel estimation method in order to overcome the pilot contamination over the uplink process for the massive multiple input multiple output (MIMO) systems for the exploitation of the low-rank property in a physical environment. In this scheme, the proposed parallel factor called PARAFAC, which derives massive MIMO systems parameters of the fading co-efficients, the direction of arrivals and delay characteristics that are caused by scattering in the environment. After the estimation of PARAFAC values, they formulate the estimation of optimization problems occurring in the three variables and get an optimal solution through the ALS algorithm. The major limitation in this proposed scheme is achieving high estimation accuracy by the AS is not possible.

Zhao et al. [8] have proposed a novel strategy for assigning the optimized pilot to minimize pilot contamination effect in the TDD mechanism-based massive multi-input multi-output-based wireless systems. The scheme consists of Chu-based sequences with a perfect matched auto-correlate property that is used for effective design and assigns the optimal value of the pilot sequences for gathering the channel state information (CSI) at the transmitter (CSIT). They presented analytical expressions with the NMSE value-based channel estimation scheme. From that, they calculate a strategy to reduce the contamination of pilot signals. This strategy of this pilot assignment scheme is widely used for improving the effective performance measures of this wireless system, by which it can attain optimal solutions. Thus, the improvised system can be used to enable reducing pilot contamination. The main disadvantage is the uplink achievable data rate versus SNR performance of TDD-based massive MIMO needs to be improved.

Muamer Hawej, et al., proposed this scheme in order to reduce the contamination of the pilots in TDD multi-cells massive MIMO system. In this scheme, this pilot contamination effect occurs only during the non-orthogonal TDD-based multiplexing over the pilot contamination over the users. This scheme comprises two estimation methods that are given on the low-rank matrix value-based approximation techniques [9]. The channel estimation is constituted of the new metrics called nuclear norm vector based on the optimization problem to improve the accuracy by mitigating the pilot contamination. The estimation method IWNN is used to increase the performance of the nuclear norm estimation value. This method needs to improve the present problems in pilot contamination with a small number iteration cost. Jiaming Li, et al., introduced a novel estimation technique as a hybrid pilot (TM pilot, which is time division-based multiplying of the pilot and TS pilot, which is time superimposed pilot) by the massive MIMO systems-based uplink channel will divide the closed form of its optimization and approximation techniques for improving the uplink achievable rate [10]. This hybrid pilot can be used with the help of both TM pilot contamination and TS pilot contamination. With the hybrid pilot, they can provide a superior range of performance to do work with TS pilot or TM pilot. The power and time ratio between pilot and data needs to be optimized.

Hayder Al-Salihi et al. have proposed a new pilot design and assigning scheme to minimize the NMSE estimators of massive MIMO system base station (BS). In this scheme, at first, the originally transmitted information signal is decomposed into convex problems into different distributed optimization problems for each and every individual BS. After this optimization, the successive approximation technique is converted into a linear matrix-based inequality form based on the optimization algorithm [11–13]. The individual BS antennas can be widely used to optimizing their own pilot data with the help of other signals from all the BS. The major disadvantage is that the proposed algorithm computational complexity is comparatively high.

3 Proposed Pilot Decontamination Algorithm

Let the multi-cell massive MIMO system model is considered with consisting of the L values of the number of hexagonal cells chosen, and every hexagonal cell widely consists of a base station with a different value of M number of antenna and K single transmitting antenna users [14]. From the j_{th} cell used in the adjacent kth user to the BS of the ith cell the channel vector $h_{(j,k),i} \in C^{M \times 1}$ will be given as,

$$h_{(j,k),i} = g_{(j,k),i}\sqrt{\beta_{(j,k),i}} \tag{1}$$

From the above expression, the large- and small-scale fading vector is represented by $\beta_{(j,k),i,}$ and $g_{(j,k),i}$, respectively. This small-scale values of the fading vector with some distribution $\mathrm{CN}(0,\mathrm{I}_M)$. At the time of the coherence period, $h_{(j,k),i}$ is the channel vector will be unaltered when it will make use of mostly used block-fading channel model. Consider that the total available $S(S \geq K)$ pilots $\varphi_i \in C^{T \times 1}$ $(1 \leq i \leq S)$ are orthogonal to each other when the length T can be used in one cell, $\phi = [\varphi_1, \varphi_2, \ldots, \varphi_S]^T \in C^{S \times T}$, $\phi\phi^H = I_{S,}$ due to limited pilot resource, the similar pilot group ϕ is reutilized [15].

By Rayleigh fading channels, the pilot reutilizes technique that can be occurring within the one cell that can be exploiting correlatively. The pilot assignment methods are used to allocate pilot $p_{(j,k)}$ to the users at the different timeslots as (j, k) non-specifically, $P(j, k)$ and that ensure that the pilot's contaminations have not to be reutilized within one of its adjacent cell $p_{(j,\,k)} \neq p_{(j,\,k')}$, The user (j, k) has the uplink SINR value for the station antennas. So, it is represented by (SINR) will be based on the matched-filter receiver is assuming to the base station. This can be solved pilot assignments as,

$$\mathrm{SINR}^{UL}_{(J,K)} = \frac{h^{H^4}_{(j,k)j}}{\sum_{(j^\circ,k^\circ)\in I_{(j,k)},\,j^\circ \neq j'} h^{H^4}_{(j,k)j} + \sigma^2_{(j,k)}/\rho^2} \tag{2}$$

$$\approx \frac{\beta^2_{(j,k),j}}{\sum_{(j_\circ,k_\circ)\in I_{(j,k)},j_\circ\neq j'}\beta^2_{(j_\circ,k_\circ),j}}\mathrm{M}\to\infty \tag{3}$$

The same pilot with a set of users has the user (*j*, *k*) and potential of uncorrelated interference is represented by $I_{(j,\,k)}$ and $\sigma^2_{(j,k)}$, respectively. The additive white Gaussian noise is always be the given essentially decreased by the rapid increasing the number of base station antennas M, $\sum_{(j',k')\in I(j,k),j_\circ\neq j'}\beta^2_{(j',k'),j}$ indicates PC that is occurring by the pilot reused ρ indicates the power for the transmission. The calculated average uplink achievable rate of the user (*j*, *k*) is,

$$C^{UL}_{(j,k)} = (1-\mu_s)E\left\{\log_2\left(1+\mathrm{SINR}^{UL}_{(J,K)}\right)\right\} \tag{4}$$

Uplink pilot transmission can because of the loss of spectral co-efficiency which can be solved by μ_s. That is the ratio of channel coherence time $\mu_\mathrm{s} = {}^{\tau}\!/_{l}$ and pilot length T. From the above equations, it is said that the small-scaling d and thermal noise are equal to fading effects. This can be calculated as M; it leads to infinity by pilot contamination.

4 Proposed WGC-Based Pilot Contamination Scheme

4.1 Problem Formulation for Optimization

The pilot assignments based for N slots for S pilot and *j*th cell is $A_s^k = S!/((S-K)!)$ for the K users. When only one cell is considered, there is no change in the A_s^k kinds of pilot assignments. *A* specific structure with L number of the hexagonal cells, the whole number of importantly various numbers of the pilot assignment is high as $(A_s^k)^{L-1}$. The allocation of the pilot $\varphi_{p(j,k)}$ to user randomly (*j*, *k*) by classical pilot assignment. In L hexagonal, the maximization of KL users can be achieved. This can be formulated as the accompanying improvement issue P_1:

$$\max_{p(j,k)}\left\{\sum_{(j,k)}\log_2\left(1+\frac{h^{H^4}_{(j,k)j}}{\sum_{(j^\circ,k^\circ)\in I_{(j,k)},j^\circ\neq j'}h^{H^4}_{(j,k)j}+\sigma^2_{(j,k)}/\rho^2}\right)\right\} \tag{5}$$

where $\{P(j,k)\}$ indicate the types of various pilot assignment and (*j*, *k*) indicate all K and L number of the users. Since the base station has not given accurate channel information. This optimization problem P_1 is impossible to solve. A huge amount of fading coefficient $\beta_{(j,k)}$ can represent the limit of total uplink throughput. The optimization problem P_1 can be approached by P_2,

$$\max_{p(j,k)}\left\{\sum_{(j,k)}\log_2\left(1+\frac{\beta^2_{(j,k),j}}{\sum_{(j_\circ,k_\circ)\in I_{(j,k)},j_\circ\neq j'}\beta^2_{(j_\circ,k_\circ),j}}\right)\right\} \tag{6}$$

The optimization problem P_2 has been solved by the complete search in $(A_s^k)^{L-1}$ kinds of the pilot assignment. The classical multi-cell massive MIMO system with various numbers of the L value is 7 and both S and K value is 8 the search difficulty is $(A_s^k)^{L-1} = (8!)^6 \approx 4.3 \times 10^{27}$.

4.2 Modified EWIG Construction

Two users in the various cells can be considered with a similar pilot, (j, k) and the inverse pilots for $P_{(j,\,k)} = P_{(j',\,k')}$, Now, modify the uplink rate of users (j, k) generated as,

$$C^{UL}_{(j,k)} \propto \log_2\left(1+\frac{\beta^2_{(j,k),j}}{\beta^2_{(j',k'),j'}+\varepsilon_{j,k,j',k'}}\right) \tag{7}$$

where $\varepsilon_{j,k,j',k'} = \sum_{(j_\circ,k_\circ)\in I_{(j,k)},j_\circ\neq j'}\beta^2_{(j',k'),j}$ denotes the pilot contamination. When the same pilot with different users can cause pilot contamination. $C^{UL}_{(j,k)}$ could be comparably represented by using pilots of the assigned users. To verify the original measurement of pilot contamination with the various $\varepsilon_{j,k,j',k'}$ of user (j, k) is very complicated.

The pilot contamination ratio is always between the user used in the vector (j', k') and users (j, k) with a similar pilot are nearly the same. The ratio of pilot contamination between the users is $\beta^2_{(j',k'),j}/\beta^2_{(j,k),j}$ and $\beta^2_{(j,k),j'}/\beta^2_{(j',k'),j'}$. The potential impact of the contamination of pilots is measured always who lies in between a large number of users in various cells, (j, k) and (j', k'),$j \neq j'$. For this operation, define a metrics is known as pilot contamination. This can be represented as,

$$\zeta_{(j,k),(j',k')} = \beta^2_{(j',k'),j}/\beta^2_{(j,k),j} + \beta^2_{(j,k),j'}/\beta^2_{(j',k'),j'} \tag{8}$$

where $\zeta_{(j,k),(j',k')}$ represents the variation of the proportion between impedance channel quality and powerful channel quality, and larger ζ represents that increasingly serious pilot contamination will be established between ore number of the user (j, k) and user over the multi-cell user (j', k') when the similar pilot is always an issue. Unmistakably pilot contamination metric as equal property $\zeta_{(j,k),(j',k')} = \zeta_{(j,k),(j',k')}$, $\forall j \neq j'$. Thus, modified EWIG has widely used be interspersed as a directionless weighted graph $G = (V, E)$, where $V = \{(j, k)\text{: } \leq j \leq L,\ 1 \leq k \leq K\}$, and $E = \{\zeta_{(j,k),(j',k')}\text{: } j \neq j'\}$, thus the V represent the vertices in all users and E represent the edges in all users have potential pilot contamination. When the edges with removable weight lies between two users in various cells are negligible for resemblance.

4.3 Modified EWIG and WGC-Based Pilot Contamination Scheme

There are many users in various cells that are placed in the EWIG. This type of assigning a similar pilot is assigned it is clear that it has the large pilot contamination $\zeta_{(j,k),(j\prime,k\prime)}$. This represents to introduce more number of pilot contaminations. When the pilot contamination $\zeta_{(j,k),(j\prime,k\prime)}$ is sufficiently low, then the performance loss is negligible for the different users in various cells has repeated similar pilot.

To investigate the important pilot assignments with significantly decreased pilot contamination by manually allocating a various number of the pilots to the attached users with a bigger iterative re-weight mechanism. For this purpose, it can use EWIG construction. EWIG is an essential tool for this operation. When the pilot resource is usually limited only K $\ll$ KL represents the more number of the N number of the orthogonal pilots of the vector placed are applicable in classical multi-cell massive MIMO systems. Under the N number of the restriction of minimal pilot resources, the WGC-PD method is proposed to achieve a change between pilot overhead and the reduction of pilot contamination and substantially reduces pilot contamination.

On the base of the typical DSATUR method which will always sort the value of the vertices range in more descending order based on its degrees and colors them as sequential based as possible with reused colors. The EWIG users are linked with the proposed WGC-PD method manually allocated various pilots with a broad weight. Unlike the DSATUR method which ensures that unused vertices are allocated with similar color, there are two users in the separate adjacent cells with a minimum weight may be allocated to the similar contaminated number of the pilots to the more and new restriction of finite pilot resources. This proposed weighted graph coloring-based pilot decontamination (WGC-PD) method will be always considered for the different variants based on the standard DSATUR method edge-weighted graph over the constrained strategies. The proposed WGC-PD involves three important main steps namely, initialization, user selection, and pilot assignment.

4.3.1 Initialization

User (j_1, k_1), and user (j_2, k_2) are two users. Such two users are selected in various cells of the EWIG. The two users are assigned their pilots φ_1 and φ_2, respectively; they are added to assigned pilots and prepare for the initialization. After that, the rest of the users are picked and sequentially allocated to all the users.

4.3.2 User Selection

An important parameter $\delta_{(j,k)}$ for select users in order of specification is introduced, which is defined as users in various cells within Ω and the weighted aggregate of the edges interface user (j, k). The user (j_0, k_0) with the largest effective pilot contamination intensity from allocation set Ω will determine whose pilot assignments must be treated favorably.

4.3.3 Pilot Assignment

After the selection of the user (j_0, k_0), the proposed reweighted-based WGC-PD method focused to pick the pilot which causes the lower efficient pilot contamination to their users from the present pilot resources. The alternative pilot set A is initially developed, which consists of the free pilots in the j_0th cell to make sure that none pilot is repeated within a similar cell $p_{(j,k)}$.

By assuming that user (j_0, k_0) is assigned to pilot S and also defined η_s to describe the effective pilot contamination potential between the adjacent cell users with various assigned pilot S in the range and user (j_0, k_0) . The pilot has the lowest effective pilot contamination potential η_s is pick to be allocated to the user (j_0, k_0) and user (j_0, k_0) will be always added into the allocated within the number of the set Ω. The process will be done in a linear manner with all users being allocated to their respective pilots.

5 Simulated Results and Discussion

The chapter gives a detailed description of the performance of the proposed modified weighted graph coloring-based pilot decontamination algorithm is investigated by using Monte Carlo simulations. In this massive MIMO system, the typical hexagonal type of cellular structures with L number of the cells is considered, where each cell has K number of single-antenna users and base station (BS) with M number of receiving and transmitting antennas. For the spectral efficiency, loss value with S is always equal to K is set as $\mu_0 = 0.05$. So the corresponding μ_s value is calculated as $\mu_s = \left(\frac{S}{K}\right) \mu_0$. The parameters for the simulation are tabulated in Table 1. The co-efficient $\beta_{<j,k>,i}$ modeled as,

$$\beta_{<j,k>,i} = \frac{z_{<j,k>,i}}{\left[\frac{r_{<j,k>,i}}{R}\right]^{\infty}} \tag{9}$$

where $Z_{\langle j,k\rangle,i}$ gives the shadow fading models and it gives a log normal probability distribution is always equal to Gaussian distributed with mean value as zero and the σ_{shadow} standard deviation $r_{\langle j,k\rangle,i}$ is measured as the distance vector matrix between Kth number of the user of jth number of the cell and βS values is the ith cell and cell radius is given as R.

Table 1 Simulation parameters

Parameters	Values
Total number of adjacent cells considered (L)	3,7
Total number of base station considered (K)	8–256
Total number of end users in the each hexagonal cells (K)	4,6
Total number of orthogonal contamination pilots (s)	$K \leq S \leq KL$
Radius of communication range of each cells (R)	500 m
Power of transmission (P)	[5,30] dB
Minimum spectral efficiency loss considered	$\mu_0 = 0.05$
Minimum path loss exponent considered	3
Log normal shadow model value	8 dB

5.1 CDF Comparison

The uplink achievable rate cumulative distribution function (CDF) curve is calculated simulated by using the MATLAB with the following parameters. The system parameters considered for the simulations are tabulated below,

The proposed WGC-PD method is simulated and comparing the proposed systems with all the other existing systems. Initially, the pilot assignment scheme is assigning the various pilots to the users randomly. In this classical pilot assignment strategy, the pilots are not having any co-operations among the cells. Here, it is considered the two cases, the BS is provided with more number of transmission and receiving antennas. These two cases are $M = 32$ and $M = 256$. By considering the optimum solution of the exhaustive searches, this simulated result shows that the proposed scheme is ensured the outperforms that all the other existing classical pilot assignment strategies (Figs. 1 and 2).

The data statistics of the curves are tabulated in Table 2 and Table 3. These values form the output result that this proposed method outperforms the other existing system.

5.2 SNR Comparison

The difference between the average powers of the transmitted signal to the average power of the noise in the communication system is called SNR. The figure shows that the average uplink achievable rate per user against the power required for the transmission or the power received at the base station (Figs. 3 and 4).

The data analytics of these curves are tabulated in Table 4. These values show that the proposed work will outperform the existing pilot assignment schemes is given in Table 5.

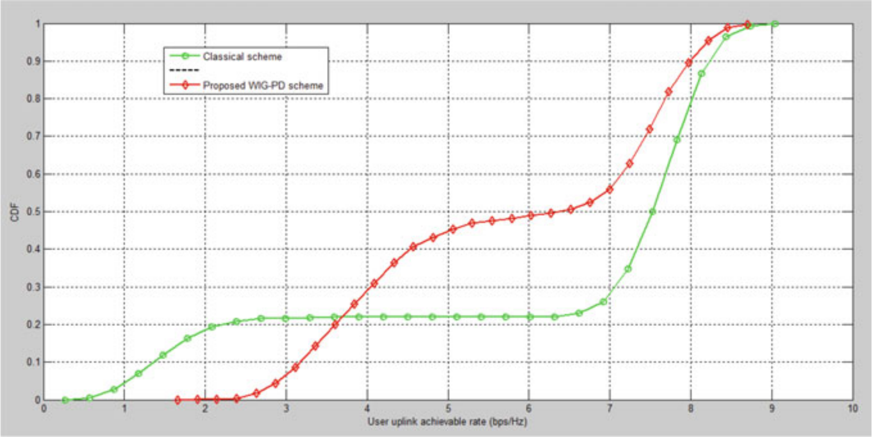


Fig. 1 CDF of the user's uplink achievable rate ($M = 32$)

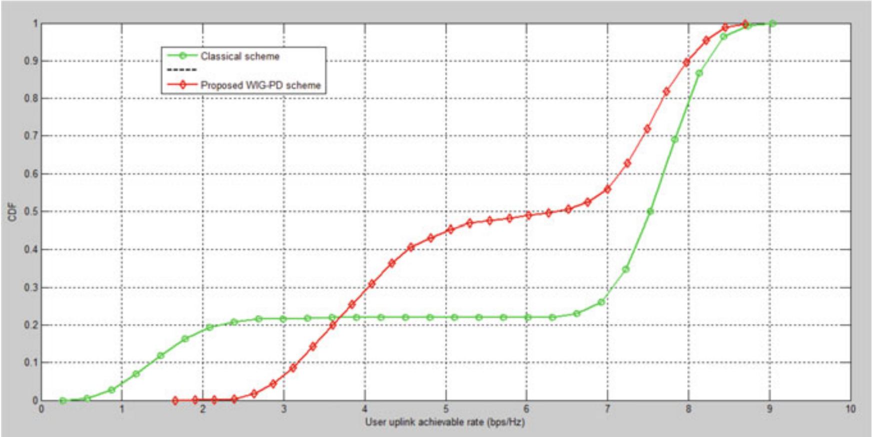


Fig. 2 CDF of the user's uplink achievable rate ($M = 256$)

In this proposed pilot scheme S and K values are the same, this will utilize the pilots effectively so this proposed scheme significantly outperforms than all other existing classical schemes. In this classical pilot assignment strategy, the pilot resources are not utilized effectively rather than assigned randomly. But in this proposed scheme, the pilots are assigned by using graph theory and weight based.

Table 2 Comparison of statistical data of CDF ($M = 32$)

Parameters	$M = 32$			
	Classical scheme		Proposed method	
	User uplink achievable rate (bps)	CDF	User uplink achievable rate (bps)	CDF
Minimum	0.2406	0	0.3365	0
Maximum	3.743	0.9975	3.479	0.9975
Mean	1.992	0.616	1.908	0.6035
Median	1.992	0.8016	1.908	0.7719
Mode	0.2406	0	0.3365	0
STD	1.063	0.3988	0.954	0.3996
Range	3.502	0.9975	3.143	0.9975

Table 3 Comparison of statistical data of CDF ($M = 256$)

Parameters	$M = 256$			
	Classical scheme		Proposed method	
	User uplink achievable rate (bps)	CDF	User uplink achievable rate (bps)	CDF
Minimum	0.2686	0	1.659	0
Maximum	9.042	0.9987	8.7	0.9975
Mean	4.655	0.3161	5.18	0.4239
Median	4.655	0.2194	5.18	0.4609
Mode	0.2686	0.2194	1.659	0
STD	2.663	0.287	2.137	0.311
Range	8.774	0.9987	7.041	0.9975

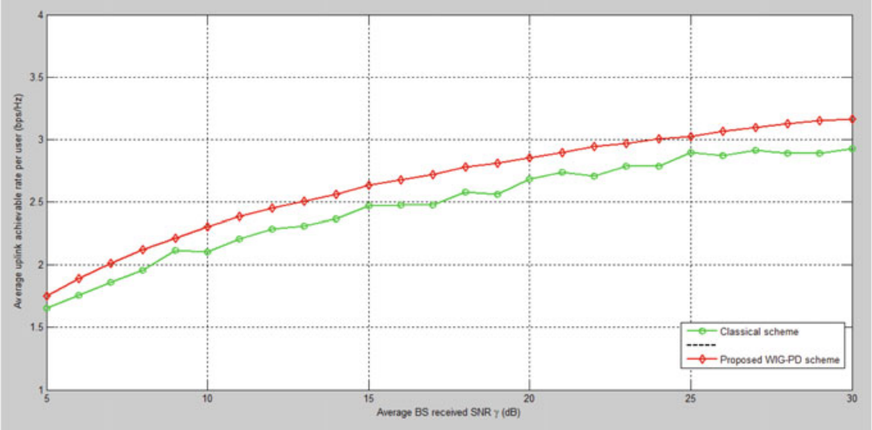


Fig. 3 Average uplink achievable rate for $M = 32$

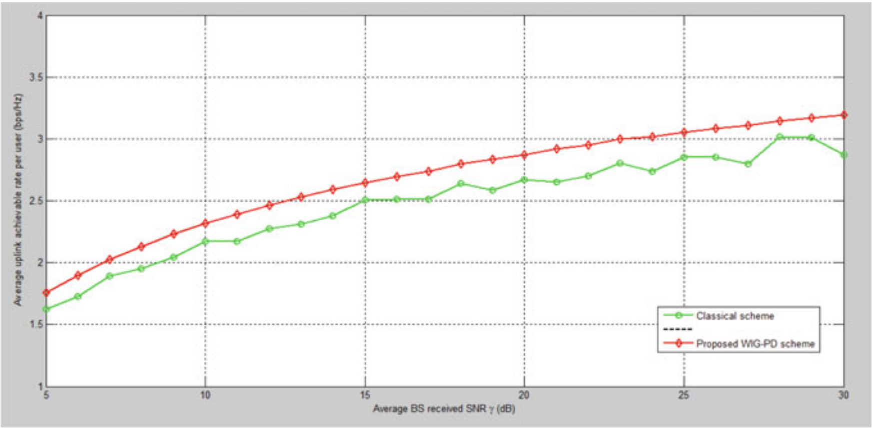


Fig. 4 Average uplink achievable rate for $M = 256$

Table 4 Comparison of statistical data of SNR ($M = 32$)

Parameters	Average uplink achievable rate ($M = 32$)		
	Average BS received SNR (dB)	Classical scheme	Proposed scheme
Minimum	5	1.652	1.75
Maximum	30	2.925	3.166
Mean	17.5	2.472	2.558
Median	17.5	2.52	2.749
Mode	5	1.652	1.75
STD	7.649	0.3845	0.4116
Range	25	1.273	1.416

Table 5 Comparison of statistical data of SNR ($M = 256$)

Parameters	Average uplink achievable rate ($M = 256$)		
	Average BS received SNR (dB)	Classical scheme	Proposed scheme
Minimum	5	1.621	1.755
Maximum	30	3.016	3.192
Mean	17.5	2.472	2.675
Median	17.5	2.548	2.766
Mode	5	1.621	1.755
STD	7.649	0.3903	0.4156
Range	25	1.394	1.437

6 Conclusion

The proposed work investigates the existing pilot assignment methods. The weighted graph coloring-based pilot decontamination assignment scheme is proposed for massive MIMO systems with 5G wireless systems to overcome and mitigate the contamination of the pilots in the assignment related issues. In this scheme, EWIG metrics are first constructed based on the available potential contamination of the PC relationships among the adjacent cell users. After the construction of EWIG, PC assigned different pilot contamination to the users in which it having the large weight in iteratively reweighted EWIG. This will reduce the unwanted pilot contains and ensures the efficient utilization of the available resources. The simulation result shows that the proposed WGC-PD scheme outperforms all the other existing systems.

References

1. Larsson EG, Edfors O, Tufvesson F, Marzetta TL (2014) Massive MIMO for next generation wireless systems. IEEE Commun Mag 52(2)
2. Rusek F et al (2013) Scaling up MIMO: opportunities and challenges with very large arrays. IEEE Sig Process Mag 30(1):40–60
3. Dai W, Milen kovic O (2009) Subspace pursuit for compressive sensing signal reconstruction. IEEE Trans Inf Theory 55:2230–2249
4. Ciuonzo D, Salvo Rossi P, Dey S (2015) Massive MIMO channel-aware decision fusion. IEEE Trans Sig Process 63(3):604–619
5. Zhu X, Wang Z, Dai L, Qian C (2015) Smart pilot assignment for massive MIMO. IEEE Commun Lett 19(9):1644–1647
6. You L et al (2015) Pilot reuse for massive MIMO transmission over spatially correlated Rayleigh fading channels. IEEE Transm Wirel Commun 14(6):3352–3366
7. Wei X, Peng W, Chen D, Ng DWK, Jiang T (2019) Joint channel parameter estimation in multi-cell massive MIMO system. IEEE Trans Commun 67(5)
8. Zhao J, Ni S, Gong Y, Zhang Q (2019) Pilot contamination reduction in TDD-based massive MIMO systems. IET Commun
9. Hawej M, Shayan YR (2019) Pilot decontamination in massive multiuser MIMO systems based on low-rank matrix approximation. IET Commun
10. Li J, Yuen C, Li D, Wu X, Zhang H (2019) On hybrid pilot for channel estimation in massive MIMO uplink. IEEE Trans Veh Technol 68(7)
11. Al-Salihi H, Van Chien T, Le TA, Nakhai MR (2018) A successive optimization approach to pilot design for multi-cell massive MIMO systems. IEEE Commun Lett 22(5)
12. Hassan N, Fernando X (2017) Massive MIMO wireless networks: an overview. MDPI J Electron Spe Issue Smart Antennas MIMO Commun 6:63
13. Al-Juboori S, Fernando X (2018) Multiantenna spectrum sensing over correlated Nakagami-m channels with MRC and EGC diversity receptions. IEEE Trans Veh Tech 67(3):2155–2164
14. Biswas A, Gupta VR (2020) Design aspects of 5G: frequency allocation, services and MIMO antennas. Eng Appl Sci Res 47(1):103–110
15. Albreem MA, Juntti M, Shahabuddin S (2019) Massive MIMO detection techniques: a survey. IEEE Commun Surv Tutorials 21(4):3109–3132

Synthesis Approach for Emotion Recognition from Cepstral and Pitch Coefficients Using Machine Learning

S. Vaijayanthi and J. Arunnehru

Abstract Emotion Recognition is a significant research domain for speech emotion identification, which includes various emotions like happy, calm, sad, angry, anxious, depressed, fearful, etc. Hence, speech emotion analysis is gaining more admiration due to the upcoming challenges in Computer Vision. Speech commands are remaining as the most prominent way for expressing ourselves as humans. There are many new methods to analyze vocal emotion synthesis using machine learning methods. This paper proposes a synthesis approach to combine the Mel Frequency Cepstral Coefficients (MFCC) with the vibration rate(PITCH) in order to characterize the emotion according to its respective vocal speech signals. The RAVDESS dataset is utilized here and the extracted features are modelled using the K-Nearest Negibhour and Decision Tree classifier for recognizing the eight emotions. The experimental results, show the efficacy of the proposed method with an overall mean accuracy rate of 87.12% for K-NN and 77.39% for Decision Tree, which outperforms the state-of-the-art results.

Keywords Speech emotion · Mel frequency cepstral coefficients · Pitch · Feature extraction · Machine learning

1 Introduction

Speech Emotion Recognition (SER) is a knowledge that extracts sensitive features from speech signals by machine. The variations and interpretations of the specific parameters and the emotional change also acquired. Almost many proposed systems combine two processing steps. The initial step is to separate the input wavelets and

---

S. Vaijayanthi (✉) · J. Arunnehru
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani Campus, Chennai, Tamilnadu, India
e-mail: vaijayanthisekar@gmail.com

J. Arunnehru
e-mail: arunnehru.aucse@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_39

extract specific features (parameters) from it. The feature extraction [1] usually means it includes a significant knowledge compression. The next step performs a grouping of similar audio files based on the extracted features. Feature extraction is a necessary part of speech emotion recognition, and it mainly associates in SER problems; hence this research introduces a novel approach for feature extraction, using MFCC to extract emotional features in dialogue signal automatically. SER is an emerging research topic in artificial intelligence, artificial therapy and pattern recognition. The broad application of this analysis is in security fields, auto supervision in healthcare areas, interactive teaching, human-computer interaction, entertainment, and so on. Speech emotion processing and recognition systems comprise of three broad sections; they are speech signal retrieval, feature extraction, and emotion recognition. A significant difficulty in this area is the automated classification of audio files. The MFCC + PITCH technique helps in extracting features from raw audio files and generates a solid description of the content. The objective is to discover the most suitable sequence of emotions, relevant to the average and vigorous-intensity of the audio file. The features are extracted based on eight different emotions.

1.1 Outline of the Work

This research paper deals with speech-based emotion recognition, which aims to identify emotions from the speech signal. The evaluation of proposed work is carried out by using RAVADESS dataset with the statement 1 is “Kids are talking by the door”, and statement 2 is “Dogs are sitting by the door”. The feature extraction techniques involve MFCC + PITCH to obtain the audio features from human speech. K-Nearest Neighbor and Decision trees classifiers are helpful in feature extraction and undergo training and testing with a 5-fold cross-validation approach.

The rest of the paper is structured as follows. Section 2 summarizes the related work. Section 3 gives an overview of the proposed technique of feature extraction. Section 4 illustrates the K-Nearest Neighbor and Decision tree classifier. Section 5 describes the datasets used, and the experimental setup Sects. 6 and 7 subsequently reports the evaluation outcomes and comparative study with various state-of-the-art models. At the end, Sect. 8 concludes the paper.

2 Prior Research

The speech data consists of emotion recognition [2], which is essential to extract the features precisely and to signify the emotional phase of different speech waves. The prime task here is to mine proficient features for the most excellent categorization of Emotions. The prior methods include the category of synthesis and analysis of speech emotion [3].

Mel Frequency Cepstral Coefficients (MFCCs) have undergone many strategies in the area of speaker identification [4] and emotional speech recognition. Comparison of Existing studies proved that MFCCs has a better way to analyze emotions with other traditional speech features. (e.g., linear predictive coding, loudness, PLP, etc.) [5]. Hansen and Bou-Ghazale [6] verified that the features built on cepstral based analysis, overtake the Linear Predictive coding (LPC) in refining the stressed speech emotion in SUSAS database. Liu [7] revealed that there is an average rise in the accuracy of 3.6% in the Gammatone Frequency Cepstral Coefficients (GFCCs) featureset aganist MFCCs for identifying emotion. Besides, different vocal prosodic and spectral features such as shimmer and glottal parameter, etc. Are also interconnected with speech emotion [8]. Liu et al. [9] interconnected voice quality parameters such as shimmer and jitter with MFCCs features in the SUSAS database to identify emotions. He proposed, a feature selection algorithm, created by using the Fisher Correlation Coefficient and correlation analysis. He classifies the Chinese speech database by using the Extreme Learning Machine (ELM) decision trees from the Institute of Automation of the Chinese Academy of Sciences (CASIA). The Fisher Criterion helps in removing redundancy features of similar audio sources extracted from the speech emotion recognition.

Recent times, a variety of different types of features has majorly used for the identification of voice-based emotion recognition. Pan et al. [10] unveiled the mixture of Pitch, Mel-energy spectrum dynamic coefficients (MEDCs), LPCC with MFCCs in SVM classifier with the help of Berlin EmoDB and Chinese emotional database (SJTU). The main difference between MFCCs and MEDC is that MFCCs requires log after the filter bank, while MEDC needs a filter bank with a logarithmic average of energies. Chen et al. [11] pulled out the energy, Pitch, Zero crossing rate (ZCR), Correlation density, Spectrum centroid and cut off frequency with five level Melfrequency energy bands.He used fractal dimension from the three-level speech recognition model which in turn helps in resolving the difficulty of speaker-independent speech emotion recognition model. The three-level categories include, pairwise six vocal emotions, representing the input as Fisher rate at each level and providing a better Classification than the previous one.

Schuller et al. [12] combined linguistic and acoustic information of emotional features with multiple-stage classifier in SVM over seven different emotional classes. The Emotional key-phrases are spotted by Belief Network using phrase-spotting. The performance comparison with a variety of classifiers include Nearest Neighbors algorithm, Support Vector Machines (SVM), Linear Classfiers and Gaussian Mixture Model (GMM) is stated with FERMUS III emotional corpus. Rao et al. [13] deliberated his work in recognizing emotions by fusing the acoustic features and facial expressions from video-based speech signal in the real-life emotional database and gives the improved performance in the Hidden Markov Model (HMM), SVM, and GMM.

Fahad et al. [14] enhanced DNN-HMM adaptive model for identifying four emotions with epochs features based on the strength of excitation (SoE), change of phase, and instantaneous pitch is combined with MFCCs using IITKGP-SEHSC and IEMOCAP databases. The primary purpose is to extract and utilize the speech features and to avoid rapid change in the vibrations of the vocal cords of the speaker. In general prosodic features, these changes are not captured, and hence the speech signal assumed to be static. The significant challenges in the prior work are limited to 4–5 emotions since many databases do not prefer a different emotions. The pre-processing steps that effectively refine data and improve the accuracy of the classifier are missing in the context of previous SER techniques. This paper proposes an efficient SER technique to process the emotions with eight different intensity labels in the vocal speech tract, to recognize emotions in real-time.

3 Feature Extraction

The extraction of individual person emotion is the most prominent part of human speech emotion recognition [15] which represents the most significant facts that are essential for future study, the following sections ensembles the representation of the feature extraction methodology with eight different input emotional states, such as happy, sad, angry, calm, disgust, neutral and surprised, are used for two dissimilar statements in this paper. Figure 1 provides an overview of the proposed work.

3.1 Mel-Frequency Cepstral Coefficients (MFCC)

The most traditional and robust audio feature extraction methodologies are the Mel-frequency cepstral coefficients (MFCC), which have 13 features. In that, eight dif-

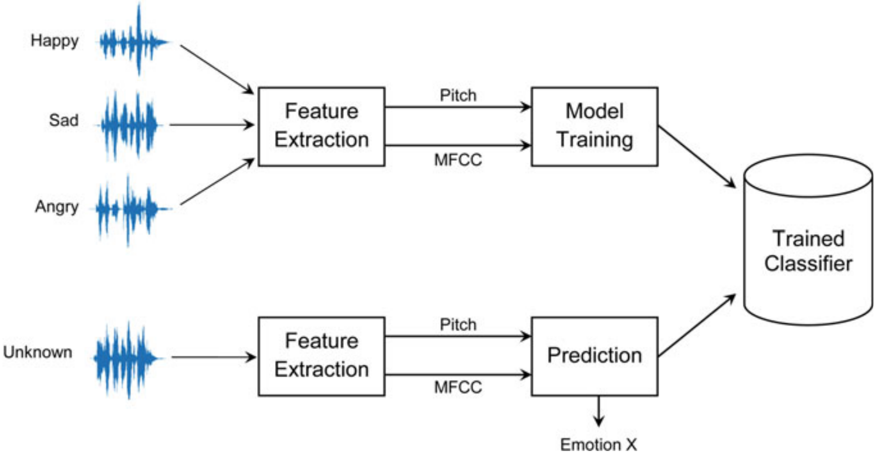


Fig. 1 Overview of the proposed work

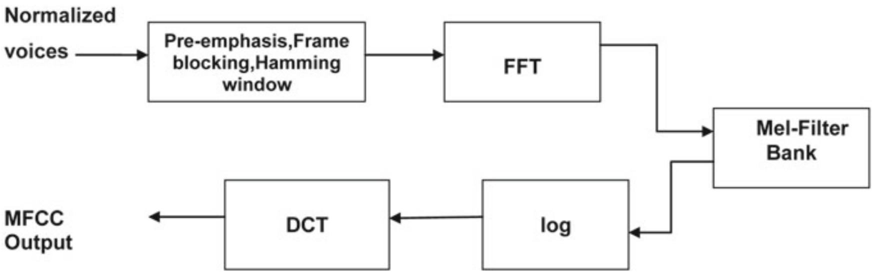


Fig. 2 MFCC feature extraction process

ferent emotional speeches are considered for classification. The feature sum is inadequate to force us to train the data of the audio. The Eight emotions were sad, happy, calm, neutral, angry, disgust, fearful and surprised etc. with two different statements. MFCC focus on interpretations of human sound perceptions, where it cannot recognize frequencies above 1 Khz [16]. It mainly involves the process of windowing the speech signal by doing triangular overlapping of the windows to take Fourier transform and map the spectrum powers to the mel scale. The third step is to find the log of the magnitude, and apply Discrete Cosine Transform to the list of mel log power, resulting is MFCC amplitudes. The complete outline of the MFCC features and the extraction steps is shown in Fig. 2.

3.1.1 Pre-emphasis, Frame Blocking, and Windowing

Pre-emphasis involves boosting the energy level and filtering to higher frequencies. The idea behind this is to poise the spectrum of vocal sounds like vowels that have a sudden roll-off in the high-frequency Zone compared with lower frequency energy levels. Hence it is called a spectral tilt, which relates to the glottal source. The glottal source in turn has −12 dB/octave drop for vocal speech signals [17]. The radiation of the sound (acoustic) energy level from the lips creates an approximately +6 dB/octave boost to the log spectrum. This results in the improvement of phone detection accuracy. On the other hand, when a vocal speech signals recorded with the help of the microphone from an average position have a downward 6 dB/octave curve corresponding to the actual spectrum of the vocal region. As a result, pre-emphasis eradicates some of the vocal tracts parameters in the glottal effects. The succeeding transfer function specifies the most frequently used pre-emphasis filter by using Eq. 1.

$$H(Z) = 1 - bz^{-1} \tag{1}$$

Here the rate of b has control over the slope of the filter between 0.4 and 1.0 [17].

The most common approaches in speech signal processing mainly relate to short-time analysis [18]. The vocal signal is a moderate quasi-stationary or time-varying signal. Regular examination of speech signals, based on steady acoustic characteristics, are examined sufficiently for a short period. The speech analysis is performed in short segments across the assumption of the speech signal is expected to be stationary. The pre-emphasis signal develops the frames of N samples. The high-frequency temporal characteristics formants process and it relates to the amplitude compared to a low frequency to attain a similar amplitude for all the formants. In all the frame, a window can minimize the speech signal with respect to frame borders. Typically, Hamming or Hanning windows are most probably used [17] to improve edge smoothing and it also decreases the edge impact of Discrete Fourier Transform (DFT) signal.

3.1.2 Fast Fourier Transform

In each frame, the Fast Fourier Transform (FFT) is imposed after windowing to discover the power spectrum of the entire frame structure. Here filter bank processing is passed out on the power spectrum, using mel-scale. Finally the Discrete Cosine Transform helpful in the conversion of windowed frames into the magnitude spectrum by using Eq. 2.

$$x\left(k\right) = \sum_{n=0}^{N-1} x(n)\ {}^{-}\,e^{\frac{j2\pi nk}{N}} \tag{2}$$

Here $x(k)$ is the frequency domain samples, $x(n)$ is the time domain samples, N exhibits the FFT size, and k ranges from ($k = 0, 1, \ldots, N - 1$).

3.1.3 Mel Scale Filter Bank

The Fourier transform moves through a series of band-pass filters to compute the Mel spectrum range known as Mel–Filter bank. Mel is a measurement of human ears received frequency in units. The Mel scale is one of the major types of linear frequency ranging below 1 khz and has a log value above 1 khz [4]. Mel estimation of physical frequency is as follows

$$f_{\text{Mel}} = 2595\log_{10}\left(1 + \frac{f}{700}\right) \tag{3}$$

In Eq. 3 (f) represents the physical frequency in Hertz, and (f_{Mel}) indicates the perceived frequency [19]. In MFCC, Filter bank's implementation carried out using frequency domain. In the frequency domain axis, the centre bank filters spaced uniformly, to identify the human ear's perception.

3.1.4 Discrete Cosine Transform

Discrete Cosine Transform is performed in MFCC to generate a pair of cepstral coefficients because the verbal stretch is smooth, and the energy level tends to be correlated. The time domain signal of the Mel log Spectrum is identified and hence the result of high frequency and low-frequency pitch signal is easily distinguished. The MFCC separates the signal coefficient information at the beginning itself, so it is easy to extract and truncate the higher-order DCT derivatives. MFCC measurement is computed in Eq. 4 as

$$C\left(n\right)=\sum_{m=0}^{M-1}\log_{10}\left(s\left(m\right)\right)\cos\left(\frac{\pi n\left(m-0.5\right)}{M}\right) \tag{4}$$

where n ranges from $n=0,1,2,\ldots,C-1$, C is the number of MFCCs, $C(n)$ represents the cepstral coefficients.

3.2 Pitch

Pitch features comply with the frequency information of the signals from our ear's response [20]. In vocal speech, the passage of air flow from the lungs is normally regulated by vocal cords, resulting in quasi-periodic excitment. A low-frequency oscillation majorly rules the subsequent sound at the end referred to as a pitch. In the non-vocal speech, the air in the lungs shrinks in the vocal tract and turns into a turbulent, noisey arousal. Depending upon the frequency of the sound wave, it is predictable as male or female. It is an essential feature of emotion recognition. Usually, the female pitch have a high pitched wave, then the males. The gender of the speaker is analyzable easily from the signal pitch by using Eq. 5. Pitch uses the auto-correlation method [21].

$$S(n)\simeq\sum_{t=x_1}^{x_2-n}y(l).y(l+n) \tag{5}$$

$S(n)$ implies the correlation of the signal n, l is the index of the signal, and y represents the signal source. Here x_1 and x_2 refer to the frame boundaries.

4 Machine Learning Algorithms

Machine learning is a subset of artificial intelligence which enables a system to attain the knowledge from data rather than through explicit programming. However, machine learning is a complex process for large dataset. As there are many strong algorithms for training data in machine learning, it is then possible to produce more precise models based on that data. This work has used K-NN and decision tree algorithms for training and testing the cepstral and pitch features.

4.1 K-Nearest Neighbors Algorithm

K-Nearest Neighbors Algorithm is a unique form of algorithm that works on the principle of supervised learning. The KNN has a set of related samples falling in a similar class of high probability. In contrast, the general idea behind the KNN algorithm is to choose K nearest neighbors for every test sample case, to predict the sample by using learnt K nearest neighbors. KNN is the non-parametric learning algorithm were no explicit training data are required. This Lazy Learner Algorithm performs well on both classification and regression related issues. During the training phase, it stores the data set, and when new data arrives, it classifies data based on a category that is very similar to new data. Here Eq. 6 uses Euclidean distance (ED) to compute the distance between two different points.

$$\mathrm{ED}(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{6}$$

The range of K value in the K-NN algorithm is still adamant and complicated [22], For example, Lall and Sharma [23] have identified the value of k is $k \simeq \sqrt{n}$ for the training datasets of training samples greater than 100 [24].

4.2 Decision Tree

A decision tree [25] is a form of supervised learning approach,which splits the voiced data spontaneously according to the specific parameter with the help of decision, problem and with the outcomes of each decision.Ambiguity is reduced in decision making, here representation of the tree is entities, called nodes and leaves. Decision tree classifier helps in classifying multi-class classification in a dataset. It undergoes the process of attribute selection with information gain and Gini index. It practices all possible results of a decision tree and traces every route to find a conclusion.

5 Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS)

The Ryerson audio-visual database of emotional speech [26] is a speech emotion dataset in the English language. This dataset contains both emotional vocal speech and song for recognition. The dataset consists of emotions of both male and female genders like happy, sad, angry, neutral, disgust, calm, surprised and fearful etc. The propsoed work has used emotional vocal speech signals alone for the proposed feature extraction approach. The sample speech signals of the eight emotions is shown in Fig. 3.

6 Experimental Results

The experiments are done using MATLAB 2019b in Windows 10 Operating System with Intel Core i7 Processor 3.40 GHz with 16 GB build-in RAM. The eight emotions are used for speech emotion recognition system viz Happy, sad, fearful, disgust, neutral, calm, surprise and angry from RAVDESS dataset [26]. The performance of the proposed method on KNN and Decision Tree classifier assessed by using a 5-fold cross-validation approach.

The satistical metrics like Accuracy (A), Precision (P), Recall (R) and F_1-measure has opted for evaluation of performance. Where tp (positive prediction) is true positive, tn (Negative prediction) is true negative, fp (mispredicted as positive) is the false positive, fn (mispredicted as negative) is the false negative. Accuracy (A)

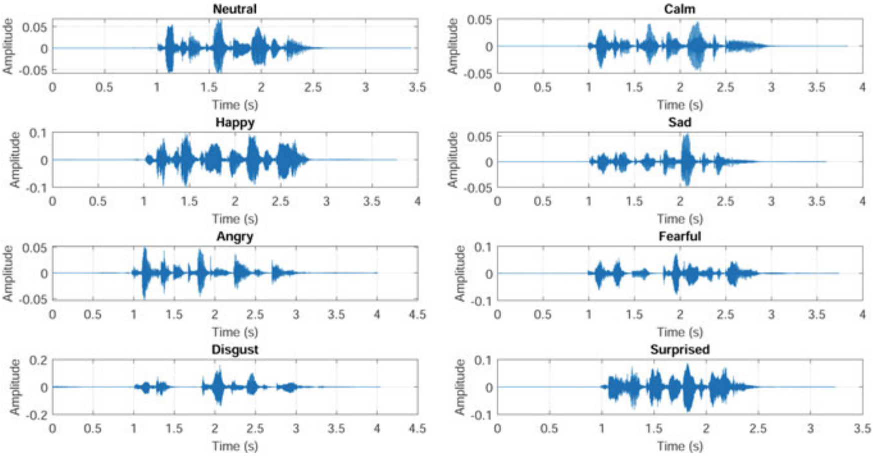


Fig. 3 Sample speech signals of the eight emotions

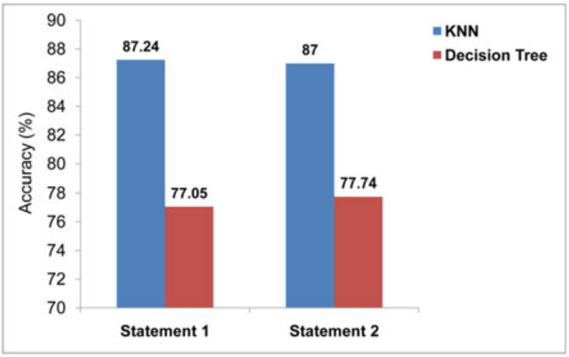


Fig. 4 The performance of K-NN and decision tree classifiers for the two statements “Kids are talking by the door” and “Dogs are sitting by the door” speech

$= \left[\frac{\text{tp+tn}}{\text{tp+fp+tn+fn}}\right]$ ensembles the overall correctness of the emotion recognition system. Precision $(P) = \left[\frac{tp}{tp+fp}\right]$ is the measure of perfection. Identification of emotions in the correct way defined by Recall $(R) = \left[\frac{\text{tp}}{\text{tp+fn}}\right]$ F_1-measure $= 2\frac{P\times R}{P+R}$ gives the mean of both Precision and Recall.

The proposed synthesis method obtained an overall accuracy rate for statements “Kids are talking by the door” and “Dogs are sitting by the door” are 87.24 and 87% for K-NN classifier and the accuracy results for Decision tree are 77.74% and 77.05% respectively. From the experimental results K-NN classifier performs well on the two speech statements when compared to decision tree. The performance graph of the KNN and decision tree for the two statements is shown in Fig. 4.

6.1 *Results on K-NN Classifier*

This section discusses the K-NN [27] and decision tree classifier performance in terms of Accuracy (A), Precision (P), Recall (R), and F_1-measure for the two vocal speech statements. The K-NN and Decision tree performance measures are shown in Tables 1 and 2. From the results, K-NN performs well when compared to decision tree and the corresponding confusion matrix for the eight different emotions using the two statements “Kids are talking by the door” and “Dogs are sitting by the door” are shown in Figs. 5 and 6 respectively.

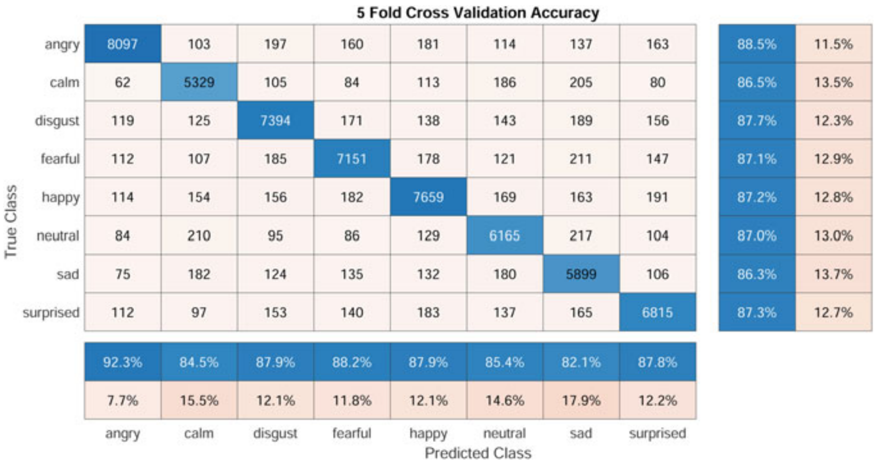
The main diagonal of the confusion matrix represents the number of instance/samples that was classified precisely/correctly. Rows represent the emotional class instance, and column represents the speech emotion class predicted by the KNN classifier. The emotions like Happy, Sad, Fear, Disgust, Neutral, Calm, surprise and Angry are classified with greater accuracy. On average, the emotion recognition rate of K-NN classifier in RAVDESS dataset is 87.12%. Here, some of the emotions like neutral and calm are misclassified as sad, since it is tough to distinguish the emotions, and it needs further attention.

Table 1 Performance measures obtained for the KNN classifier

Emotion	"Kids are talking by the door"				"Dogs are sitting by the door"			
	Precision	Recall	Specificity	*F*-score	Precision	Recall	Specificity	*F*-core
Angry	0.9227	0.8847	0.9873	0.9033	0.9123	0.8883	0.9847	0.9002
Calm	0.8449	0.8645	0.9826	0.8546	0.8419	0.8723	0.9794	0.8568
Disgust	0.8793	0.8766	0.9812	0.8779	0.8756	0.8731	0.9823	0.8743
Fearful	0.8819	0.8708	0.9824	0.8763	0.8697	0.8738	0.9807	0.8717
Happy	0.8790	0.8715	0.9804	0.8753	0.8767	0.8697	0.9820	0.8732
Neutral	0.8545	0.8695	0.9810	0.8619	0.8451	0.8503	0.9784	0.8477
Sad	0.8209	0.8633	0.9769	0.8416	0.8369	0.8534	0.9802	0.8451
Surprised	0.8780	0.8735	0.9827	0.8757	0.8877	0.8729	0.9840	0.8802
Mean (μ)	0.8702	0.8718	0.9818	0.8708	0.8682	0.8692	0.9815	0.8686

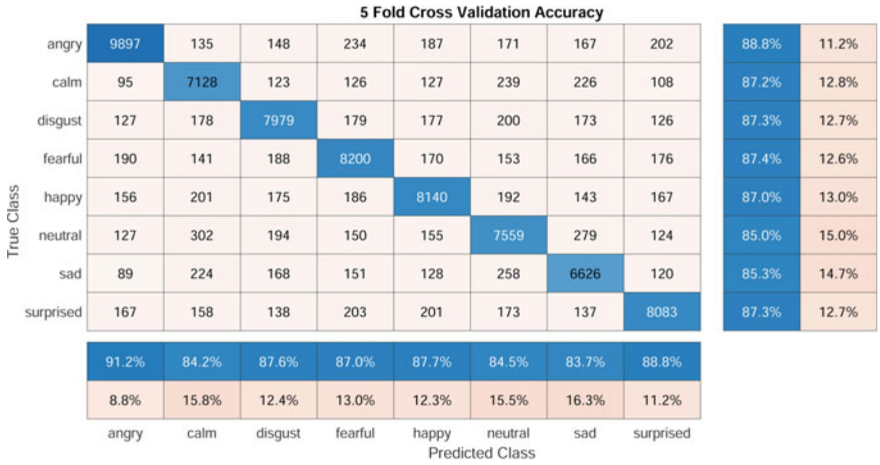
Table 2 Performance measures obtained for the Decision tree classifier

Emotion	"Kids are talking by the door"				"Dogs are sitting by the door"			
	Precision	Recall	Specificity	*F*-score	Precision	Recall	Specificity	*F*-Score
Angry	0.8368	0.7757	0.9740	0.8051	0.8311	0.7896	0.9721	0.8098
Calm	0.7502	0.7837	0.9714	0.7666	0.7570	0.7883	0.9669	0.7723
Disgust	0.7615	0.7615	0.9628	0.7615	0.7794	0.7747	0.9670	0.7770
Fearful	0.7735	0.7852	0.9652	0.7793	0.7788	0.8041	0.9672	0.7912
Happy	0.7885	0.7816	0.9657	0.7850	0.7822	0.7719	0.9693	0.7770
Neutral	0.7462	0.7618	0.9668	0.7539	0.7303	0.7545	0.9620	0.7422
Sad	0.7264	0.7676	0.9645	0.7464	0.7532	0.7739	0.9687	0.7634
Surprised	0.7658	0.7470	0.9674	0.7563	0.7983	0.7599	0.9724	0.7786
Mean (μ)	0.7686	0.7705	0.9672	0.7693	0.7763	0.7771	0.9682	0.7765



	angry	calm	disgust	fearful	happy	neutral	sad	surprised		
angry	8097	103	197	160	181	114	137	163	88.5%	11.5%
calm	62	5329	105	84	113	186	205	80	86.5%	13.5%
disgust	119	125	7394	171	138	143	189	156	87.7%	12.3%
fearful	112	107	185	7151	178	121	211	147	87.1%	12.9%
happy	114	154	156	182	7659	169	163	191	87.2%	12.8%
neutral	84	210	95	86	129	6165	217	104	87.0%	13.0%
sad	75	182	124	135	132	180	5899	106	86.3%	13.7%
surprised	112	97	153	140	183	137	165	6815	87.3%	12.7%
	92.3%	84.5%	87.9%	88.2%	87.9%	85.4%	82.1%	87.8%		
	7.7%	15.5%	12.1%	11.8%	12.1%	14.6%	17.9%	12.2%		

Fig. 5 Confusion matrix obtained for the KNN classifier for Statement 1 "Kids are talking by the door"


True Class
Predicted Class

5 Fold Cross Validation Accuracy

	angry	calm	disgust	fearful	happy	neutral	sad	surprised		
angry	9897	135	148	234	187	171	167	202	88.8%	11.2%
calm	95	7128	123	126	127	239	226	108	87.2%	12.8%
disgust	127	178	7979	179	177	200	173	126	87.3%	12.7%
fearful	190	141	188	8200	170	153	166	176	87.4%	12.6%
happy	156	201	175	186	8140	192	143	167	87.0%	13.0%
neutral	127	302	194	150	155	7559	279	124	85.0%	15.0%
sad	89	224	168	151	128	258	6626	120	85.3%	14.7%
surprised	167	158	138	203	201	173	137	8083	87.3%	12.7%
	91.2%	84.2%	87.6%	87.0%	87.7%	84.5%	83.7%	88.8%		
	8.8%	15.8%	12.4%	13.0%	12.3%	15.5%	16.3%	11.2%		

Fig. 6 Confusion matrix obtained for the KNN classifier for Statement 2 “Dogs are sitting by the door”

7 Comparative Study

The results obtained from the proposed work have compared with the state-of-the-art results to show the effectiveness of the proposed technique and comparison is presented in Table 3. Based on the study, in RAVDESS database, Zeng et al. [28] has proposed a multi-task model using GresNets Spectrogram generations, with an accuracy of 64.48%. Bhavan et al. [2] evaluated the recognition of emotion considering Bagged SVM using MFCC derivatives and Spectral centroids and resulted in the best accuracy of 75.69%. Kwon et al. [29] proposes Deep Stride CNN model using Raw and Clean Spectrograms and provides a better accuracy rate of 79.50%. It is observed that our proposed approach tends to increase the recognition accuracy of 87.12% and gives excellent results in the RAVDESS dataset.

Table 3 State-of-the-art results on RAVDESS dataset

Method	Classifier	Feature	Accuracy (%)
Zeng et al. [28]	DNNs	Spectrograms	64.48
Bhavan et al. [2]	Bagged SVMs	MFCC	75.69
Kwon et al. [29]	DSCNN	Spectrograms	79.50
Ours	KNN	Cepstral + Pitch	87.12

8 Conclusion

This research paper proposes an efficient framework for speech emotion recognition using MFCC + Pitch features with K-NN and decision algorithm. Experimental analysis has conducted on vocal speech statements “Kids are talking by the door” and “Dogs are sitting by the door” from RAVDESS emotion dataset with eight different emotions. The proposed approach based on speech emotion on different statements is extracted by using MFCC + Pitch from the vocal speech signals. The extracted features are trained and modelled using K-Nearest Neighbour and Decision tree classifiers for identifying emotions. The experimental results shows that RAVADESS dataset revealed the feasibility of the proposed method with an overall accuracy of 87.12% for K-NN and 77.39% for Decision tree. Further, it has concluded that the experiments based on K-NN performs better than decision tree. The accuracy of various Quantitative evaluations is computed with metrics like precision, recall, accuracy and *F*-measure. The findings concluded that the system could not be able to distinguish neutral and calm with high precision. Our further research extends to recognize the identification of emotional patterns by manipulating gender in the perceived speech signal.

References

1. Koduru A, Valiveti HB, Budati AK (2020) Feature extraction algorithms to improve the speech emotion recognition rate. Int J Speech Technol 23(1):45–55
2. Bhavan A, Chauhan P, Shah RR et al (2019) Bagged support vector machines for emotion recognition from speech. Knowl-Based Syst184:104886
3. Kim EH, Hyun KH, Kim SH, Kwak YK (2009) Improved emotion recognition with a novel speaker-independent feature. IEEE/ASME Trans Mechatron 14(3):317–325
4. Hasan Md R, Jamil M, Rahman MGRMS et al (2004) Speaker identification using mel frequency cepstral coefficients. Variations 1(4) (2004)
5. Dave N (2013) Feature extraction methods IPC, PLP and MFCC in speech recognition. Int J Adv Res Eng Technol 1(6):1–4
6. Bou-Ghazale SE, Hansen JHL (2000) A comparative study of traditional and newly proposed features for recognition of speech under stress. IEEE Trans speech Audio Process 8(4):429–442 (2000)
7. Liu GK (2018) Evaluating gammatone frequency cepstral coefficients with neural networks for emotion recognition from speech. arXiv:1806.09010
8. Shashidhar G, Koolagudi K, Sreenivasa R (2012) Emotion recognition from speech: a review. Int J Speech Technol 15(2):99–117
9. Liu Z-T, Min W, Cao W-H, Mao J-W, Jian-Ping X, Tan G-Z (2018) Speech emotion recognition based on feature selection and extreme learning machine decision tree. Neurocomputing 273:271–280
10. Pan Y, Shen P, Shen L (2012) Speech emotion recognition using support vector machine. Int J Smart Home 6(2):101–108
11. Chen L, Mao X, Xue Y, Cheng LL (2012) Speech emotion recognition: features and classification models. Digital Signal Process 22(6):1154–1160

12. Schuller B, Rigoll G, Lang M (2004) Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine-belief network architecture. In: 2004 IEEE international conference on acoustics, speech, and signal processing, vol 1. IEEE, pp I–577
13. Sreenivasa Rao K, Koolagudi SG (2015) Recognition of emotions from video using acoustic and facial features. Signal Image Video Process 9(5):1029–1045
14. Fahad Md, Yadav J, Pradhan G, Deepak A et al (2018) DNN-HMM based speaker adaptive emotion recognition using proposed epoch and MFCC features. arXiv:1806.00984
15. Arunnehru J, Kalaiselvi Geetha M (2017) Automatic human emotion recognition in surveillance video. In: Intelligent techniques in signal processing for multimedia security. Springer, pp 321–342
16. Muda L, Begam M, Elamvazuthi I (2010) Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques. arXiv:1003.4083
17. Picone JW (1993) Signal modeling techniques in speech recognition. Proc IEEE 81(9):1215–1247
18. Benesty J, Mohan Sondhi M, Huang Y (2007) Springer handbook of speech processing. Springer
19. Deller JR, Proakis JG, Hansen JHL (2000) Discrete-time processing of speech signals. Institute of Electrical and Electronics Engineers
20. Kurpukdee N, Kasuriya S, Chunwijitra V, Wutiwiwatchai C, Lamsrichan P (2017) A study of support vector machines for emotional speech recognition. In: 2017 8th International conference of information and communication technology for embedded systems (IC-ICTES). IEEE, pp 1–6
21. Selvaraj M, Bhuvana R, Padmaja S (2016) Human speech emotion recognition. Int J Eng Technol 8:311–323
22. Kang P, Cho S (2008) Locally linear reconstruction for instance-based learning. Pattern Recogn 41(11):3507–3518
23. Meesad P, Hengpraprohm K (2008) Combination of kNN-based feature selection and kNN based missing-value imputation of microarray data. In: 2008 3rd International conference on innovative computing information and control. IEEE, pp 341–341
24. Lall U, Sharma A (1996) A nearest neighbor bootstrap for resampling hydrologic time series. Water Resources Res 32(3):679–693
25. Badshah AM, Ahmad J, Lee MY, Baik SW (2016) Divide-and-conquer based ensemble to spot emotions in speech using MFCC and random forest. arXiv:1610.01382
26. Livingstone SR, Russo FA (2018) The Ryerson audio-visual database of emotional speech and song (RAVDESS): a dynamic, multimodal set of facial and vocal expressions in North American English. PloS One 13(5):e0196391
27. Arunnehru J, Yashwanth A Shammer S (2017) Canonical correlation-based feature fusion approach for scene classification. In: International conference on intelligent systems design and applications. Springer, pp 134–143
28. Zeng Y, Mao H, Peng D, Yi Z (2019) Spectrogram based multi-task audio classification. Multimedia Tools Applications 78(3):3705–3722
29. Kwon S et al (2020) A CNN-assisted enhanced audio signal processing for speech emotion recognition. Sensors 20(1):183

Performance Analysis of Deep CNN Assisted Optimized HIV-I Protease Cleavage Site Prediction with Hybridized Technique

Navneet Kaur and Wiqas Ghai

Abstract In recent years, human immunodeficiency virus infection and acquired immune deficiency syndrome (HIV/AIDS) has emerged as a global health issue. The disease is caused by a virus that affects the CD4 cell in the human body that lowers the immune system in the human body. HIV-protease is the agent that replicates itself and affects the CD4 T cells in the human blood. To overcome the problem of replication, inhibitors can be analyzed and designed that can bind the active sites in the proteases. To design efficient protease inhibitors, the knowledge about the specificity of cleavage sites is essential. Several encoding techniques and classifiers have been proposed to study and analyze the active cleavage sites in proteases. This paper proposes a new model and comparatively analyses the performance of Hybridized SVM_Genetic modeling with Deep CNN assisted optimized prediction of Cleavage sites. For optimal tuning of activation functions, two metaheuristic algorithms such as moth search and dragonfly are proposed in this work. The performance of both the methodologies is compared based on different parameters such as accuracy, specificity, F1 score, sensitivity, and NPV. To authenticate the performance of the proposed model, standard data from machine learning algorithms called UCI repository is processed for experimentation. The performance measured is compared with existing available techniques for predicting cleavages.

Keywords Moth search · Dragonfly · Deep convolution neural network · Support vector machine · Acquired immunodeficiency syndrome

1 Introduction

Acquired immunodeficiency syndrome (AIDS) is caused due to the presence of human immunodeficiency virus present in the body. Human immunodeficiency virus

---

N. Kaur (✉) · W. Ghai
RIMT University, Mandi Gobindgarh, Punjab, India
e-mail: Bawa.navneet@gmail.com

W. Ghai
e-mail: ghaialpha@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_40

(HIV) shows symptoms of the lower body's immune system and leading to the death. The first case of HIV-1 was reported in 1981 in the center for disease controlling the USA. It is been 39 years and still the human immunodeficiency virus (HIV-I) is a global health issue. According to the World Health Organization, 38 million people living in the world are infected with infectious disease by the end of 2019 and have claimed 33 million lives so far. However, with concerted efforts of various researchers, a new technique antiretroviral therapy is devised. This kind of treatment of HIV-I smacks and suppresses the action of HIV-I proteases [1, 2]. The first reported antagonist of HIV-1 proteases was discovered in 1987. When HIV-AIDS enters the human body and targets CD4 T cell present in the blood, and replicates itself producing its copies in the blood, which lowers the body's immune system which is shown in Fig. 1(L-R) the long viral proteins are being cut by proteases, to prevent the cleavage protease inhibitors can be developed, but the problem lies in predicting the cleavage site in various viral proteins. The various antagonist and inhibitors were proposed like Saquinavir and Nelfinavir have been proposed by FDA. The basic mechanism lies in the fact is to replace the peptide linkage consisting of –NH–CO with hydroxyl ethylene group (–CH_2–CH (OH)–) which proteases is difficult to cleave [3].

To develop a standard benchmark inhibitor for HIV-proteases, so the study of HIV-1 proteases cleavage specificity is a major concern to researchers. Proteases get attached viral protein and replicate itself to produce a large number of viral proteins.

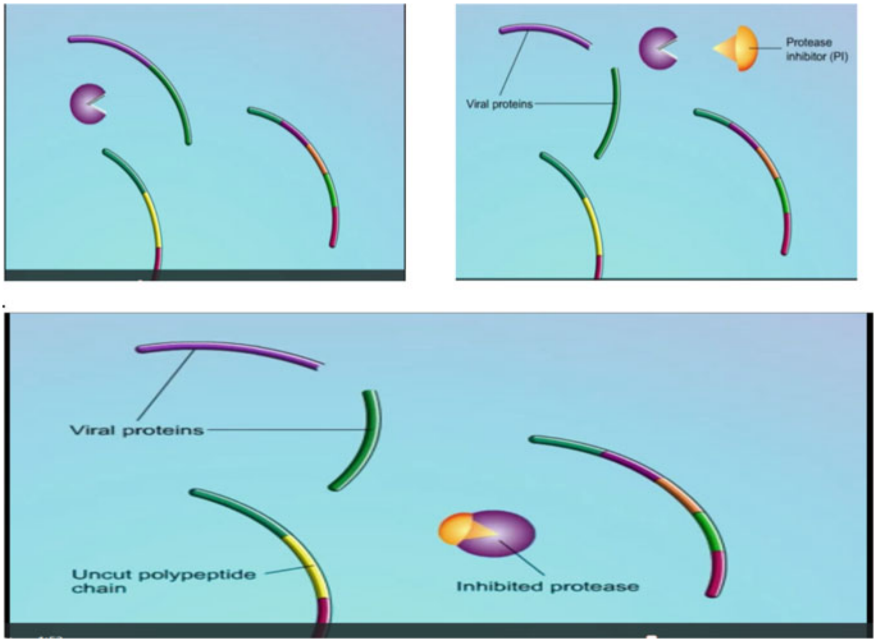


Fig. 1 (L-R) various stages of protease effecting CD4 T cells and viral protein replicating itself

2 Literature Survey

Extensive work has been done by the researchers and organizations to explore the potential of data provided in the repository for analysis, and resulted in substantial classification accuracy. Singh et al. in 2018 proposed an evolutionary-based ensemble framework, genetic algorithm to attain the optimally coupled feature and classifier sequence. The simulation showed significant enhancement of the proposed model over the conventional models because of average accuracy, sensitivity, and precision. The average accuracy achieved without cross-validation is 87.14% [4]. Fatehi and sadegi in 2018 introduced a novel technique in spatial and structural features are taken into account, SVM and genetic programming have been used for modeling, and the prediction accuracy achieved is 91.1% [5]. In 2018, Singh et al. developed an optimal formation of encoding-classifier pair selection by an evolutionary algorithm. Natural selection is based on several learners and optimal data-learner mapping. The accuracy achieved in this case was quite high [6]. In 2009, Ogul constructed a model entirely based on a generalization of variable order Markov chains (VOMC) for protein sequence; the model developed predicted cleavability by some proteases. The novel method, called variable context Markov chains (VCMC), tries to analyze the context equivalence based on the similarities of specific amino acids. The result shows that it outperforms the performance of existing models in terms of prediction accuracy on a benchmark dataset [7]. In 2016, Singh et al. proposed a prediction method in which certain properties like sequential, structural, and chemical features are added in certain machine learning methods. The main features are extracted by various encoding schemes and as input to decision trees, regression and ANN, a three-way approach was applied for prediction which achieved an accuracy of 80.0–97.4% [8]. In 2005, Lummini and Nanni constituted an idea from various machine learning algorithms to develop a knowledge base [9]. Song et al., in (2019) proposed a web server for the prediction of cleavage sites by many different proteases, using SVR with a combination of different features. Bi-profile Bayesian was used for feature extraction, and the Gini score was used for calculation. The data was used from a larger dataset from Schilling and other sources published data for cleaving whole proteins [10].

3 Material and Methods

Deep convolution neural network (DCNN) is employed for HIV-1 protease cleavage specificity prediction. It is used for classification and two metaheuristic approaches moth search and dragonfly is used for the optimization of the results that predicts the specificity, accuracy, sensitivity of the input data. The proposed model consists of following different phases such as *s*electing the dataset from the repository, data simplification, and preprocessing, feature selection model based on discrete wavelet transform, training and classification of the dataset using deep convolution neural network, and finally, optimal tuning of activation function using hybridized metaheuristic algorithms (moth search and dragonfly).

A. **Selecting Dataset from Repository**

Collecting and selection of data from the domain that is authentic, informative, and useful are very challenging, especially when you are dealing with biological data. In this study of the cleavage site, the dataset is taken from the UCI machine learning repository is a collection of databases, domain theories, and data generators that are used by the machine learning community for the empirical analysis of machine learning algorithms. The archive was created as an FTP archive in 1987 by David Aha and fellow graduate students at UC Irvine. In the proposed work, 4 datasets are used Data-746, Data-1625, data-Schilling, Data-Impens.

Each input query consists of two parts, first eight-letter alphabetic strings that represent 8 different amino acids, −1 and 1 denote the cleave and non-cleave sites in octamer. The character string that are allowed for encoding consists of {A,B,N,D,C,Q,E,H,I,L,K,M,F,P,S,T,W,Y,V} each denoting different amino acids (Table 1).

The encoding of an octamer sequence requires much attention and is necessary for interpretation by different machine learning approaches. Different encoding techniques are designed by researchers in the past like orthonormal encoding (OE), consisting of a 20-bit vector and but the main drawback is information loss. Another encoding technique includes combining BLOSUM50 and BLOSUM62 matrices with orthonormal encoding [9], Taylor Venn diagram encoding is also experimented by Zvelebil et al. The proposed encoding model in this research work consists of orthonormal encoding that combines the structural and chemical features of every amino acid. The features polarity, acidity, hydropathy index, aliphatic, aromatic, proline, etc., are taken and their values are normalized using the following formula

$$nv_i = \left(\left(\frac{r_i - \min(r_i)}{\max(r_i) - \min(r_i)} \right) \times 2 \right) \tag{1}$$

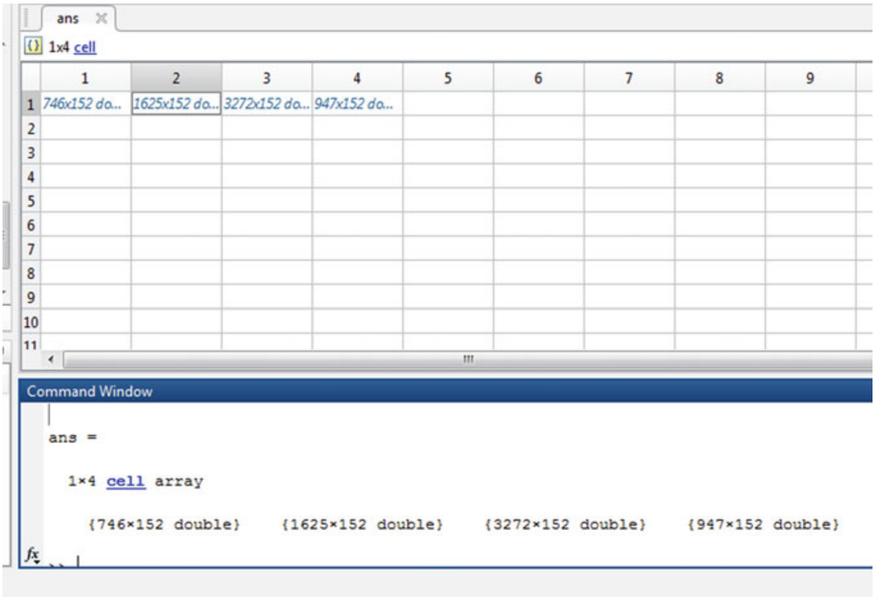
r_i and nv_i denote the original and new value in the range $1 < nv_i < 152$ (Fig. 2).

B. **Feature Selection model based on wavelet transform**

Different methodologies are applied by various researchers for feature extraction like SVM, genetic programming, ANN, KNN, etc., but the technique applied in the

Table 1 Dataset with their cleavage specificity

Sr.No	Dataset	No. of cleavages	No. of non-cleavages
1	Data_746	401	345
2	Data_1625	374	1251
3	Data_schlling	434	2838
4	Data_impens	149	798



	1	2	3	4	5	6	7	8	9
1	746x152 do...	1625x152 do...	3272x152 do...	947x152 do...					

```
ans =

  1×4 cell array

    {746×152 double}    {1625×152 double}    {3272×152 double}    {947×152 double}
```

Fig. 2 Snapshot of MATLAB showing feature preprocessing

proposed model for extracting features is wavelet-based. There are few reasons, this technique is applied including precise details of the dataset or signal that can be extracted easily, it provides a good approximation of the result and can recognize breakpoints, trends, and discontinues. Wavelet decomposition is a wavelet transform or decomposition of the signal or data into 1 dimension or 2 dimension. Suppose a signal x is divided into n levels using different wavelets. The output decomposition structure contains the wavelet decomposition vector c and the bookkeeping vector l, which maintains the coefficients by levels. The structure is shown in the decomposition diagram (Fig. 3).

Fig. 3 Representation of 1D wavelet transform for extracting features

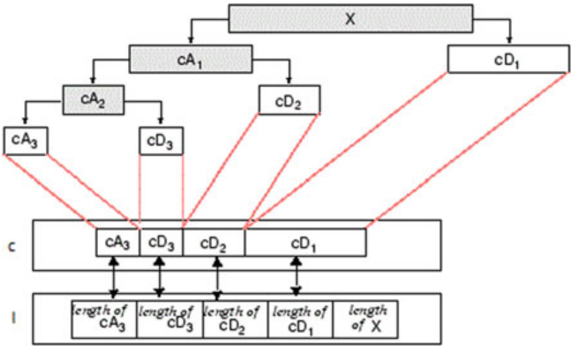
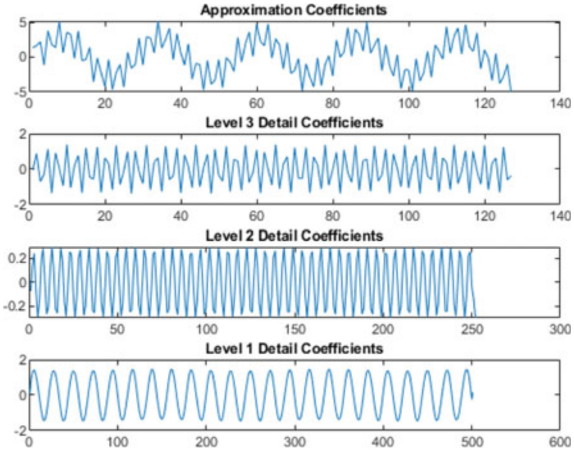


Fig. 4 Wavelet transformation after preprocessing of the data



$$[\mathrm{c}, \mathrm{l}] = \mathrm{wavedec}(x, n, \mathrm{wname}) \tag{2}$$

The wavelet decomposition equation is given by

$$[\mathrm{c}, \mathrm{l}] = \mathrm{wavedec}(x, n, \mathrm{LoD}, \mathrm{HiD}) \tag{3}$$

From Eqs. 1–3, the wavelet transformation of dataset is shown in Fig. 4.

C. Training and Classification of the dataset using deep convolution neural network

After extracting features, using wavelet transform the output extracted for the previous phase is used to train the deep convolution neural network. Classification involves a supervised learning technique, classifier learns through various attributes present in the input data, the classification is done in two steps first is training the data and then predicting the samples. The classifier used in research work is a deep convolution neural network. Classification is done by training the known sample set (UCI repository), the performance of the model is measured in terms of accuracy, which is achieved while predicting the unknown sample. CNN is divided into two parts: feature detection layers and classification layers. The pseudocode for deep CNN training is shown in Fig. 5.

D. Optimization of activation function using moth search and dragonfly algorithm

After training and optimization, the next step involves optimization of the output after the classification phase. Two metaheuristic algorithms are used to optimize the performance of deep convolution neural network. Moth search algorithm works on the phenomenon of phototaxis and Levy walk. Moths tend to fly near to the source of light, this behavior is known as phototaxis, Levy walk are one of the most important style of moths to fly in natural environments. Many animals and species,

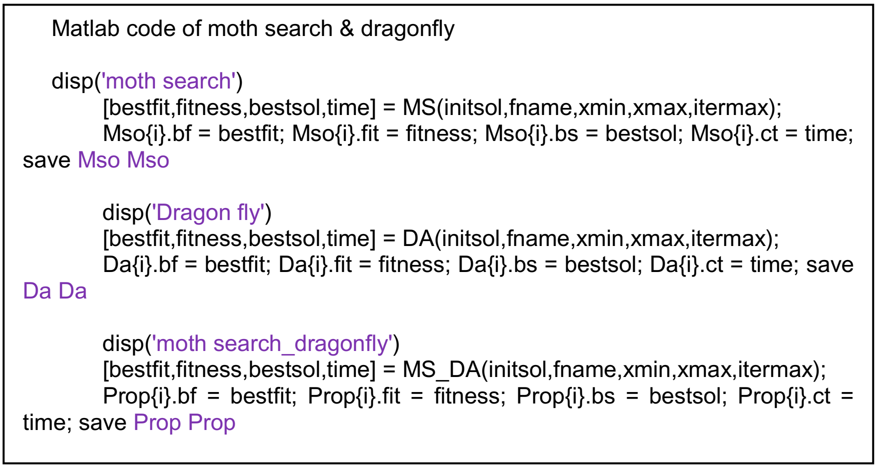
```
Matlab code for Training dataset using DCNN


function[acc,net] = Deep_CNN(train_data,train_lab,act_fn,sol)

if act_fn == 1
   a = reluLayer;
elseif act_fn == 2
   a = leakyReluLayer;
elseif act_fn == 3
   a = clippedReluLayer(10);
elseif act_fn == 4
   a = eluLayer;
elseif act_fn == 5
   a = tanhLayer;
else
   a = preluLayer;
end
layers = [
   imageInputLayer([size(train_data,1) size(train_data,2) 1])
   convolution2dLayer(sol(1),sol(2),'Padding','same')
   convolution2dLayer(sol(1),sol(2),'Padding','same')
   convolution2dLayer(sol(1),sol(2),'Padding','same')
   convolution2dLayer(sol(1),sol(2),'Padding','same')
   a
   fullyConnectedLayer(384) % 384 refers to number of neurons in next FC hidden layer
   fullyConnectedLayer(384)
   fullyConnectedLayer(length(unique(train_lab)))
   softmaxLayer
   classificationLayer
   ];
options = trainingOptions('sgdm',...
   'MaxEpochs',100, ...
   'Verbose',true);
labels = categorical(train_lab);
net = trainNetwork(train_data,labels',layers,options);
predictedLabels = classify(net,train_data)';
```

Fig. 5 Pseudocode for deep CNN

like Drosophila, fly in the form of Levy flight that can be rounded to a power law. It is basically distributed over varied scale with the feature of exponents close to 3/2 [11, 12]. It is one of the commonly used metaheuristic optimization algorithm based on swarm intelligence. The main approach used in this is based on static and dynamic behaviors of swarming in dragonflies in the environment. The two most important phases of optimization are exploration and exploitation which are configured by modeling the behavior of dragonflies, i.e., navigating, hunt for food, protecting themselves from enemies [13, 14] (Fig. 6).



```
   Matlab code of moth search & dragonfly

   disp('moth search')
         [bestfit,fitness,bestsol,time] = MS(initsol,fname,xmin,xmax,itermax);
         Mso{i}.bf = bestfit; Mso{i}.fit = fitness; Mso{i}.bs = bestsol; Mso{i}.ct = time;
save Mso Mso

         disp('Dragon fly')
         [bestfit,fitness,bestsol,time] = DA(initsol,fname,xmin,xmax,itermax);
         Da{i}.bf = bestfit; Da{i}.fit = fitness; Da{i}.bs = bestsol; Da{i}.ct = time; save
Da Da

         disp('moth search_dragonfly')
         [bestfit,fitness,bestsol,time] = MS_DA(initsol,fname,xmin,xmax,itermax);
         Prop{i}.bf = bestfit; Prop{i}.fit = fitness; Prop{i}.bs = bestsol; Prop{i}.ct =
time; save Prop Prop
```

Fig. 6 MATLAB code of moth search and dragonfly

4 Results and Discussion

The experimental setup deployed for the intel® core i5 CPU@ 1.8 GHz with 5 GB RAM, MATLAB 2018(9.4.0) is used that combines a desktop environment tuned for iterative analysis and design processes the expression matrix and array mathematics directly. Dataset is used for training and testing; i.e., it is taken from the UCI repository. The performance of the proposed CNN model is compared with the existing state-of-the-art models, work done by researchers is evaluated based on accuracy, specificity, precision, FPR, F1 score, NPV, MCC, FDR, etc.

1. Accuracy:T_p denotes correctly predicted cleaved and T_n denotes correctly predicted non-cleaved sites and F_p and F_n number of incorrectly predicted cleaved and incorrectly predicted non-cleaved sites [13, 15].

$$\text{Accuracy} = \frac{T_\mathrm{p} + T_\mathrm{n}}{T_\mathrm{p} + T_\mathrm{n} + F_\mathrm{p} + F_\mathrm{n}}$$

The accuracy achieved by various datasets is shown in Table 2.

Average accuracy achieved using a benchmark dataset in the proposed methodology, i.e., deep CNN assisted HIV-1 protease cleavage site prediction is 93%, which is quite high as compared to other states-of-the-art techniques (Fig. 7 and Tables 3, 4, 5).

2. Sensitivity: It specifies the likelihood of false positives. It is defined as the True positive rate given by the following formula. In the proposed algorithm, the average sensitivity achieved is 95.53% which is comparatively better as compared to other processes. The sensitivity is given by the formula

Table 2 Performance classifier accuracy achieved with various techniques (existing and proposed)

Technique dataset	DCNN	SVM_GA	MSO	DA	DA_MSO (proposed method)
Data746	84.375	81.696	89.286	90.625	92.411
Data1625	93.634	81.696	92.197	93.84	94.661
DataShilling	92.974	81.696	93.177	93.279	93.89
DataImpens	88.732	81.696	90.493	88.028	91.197

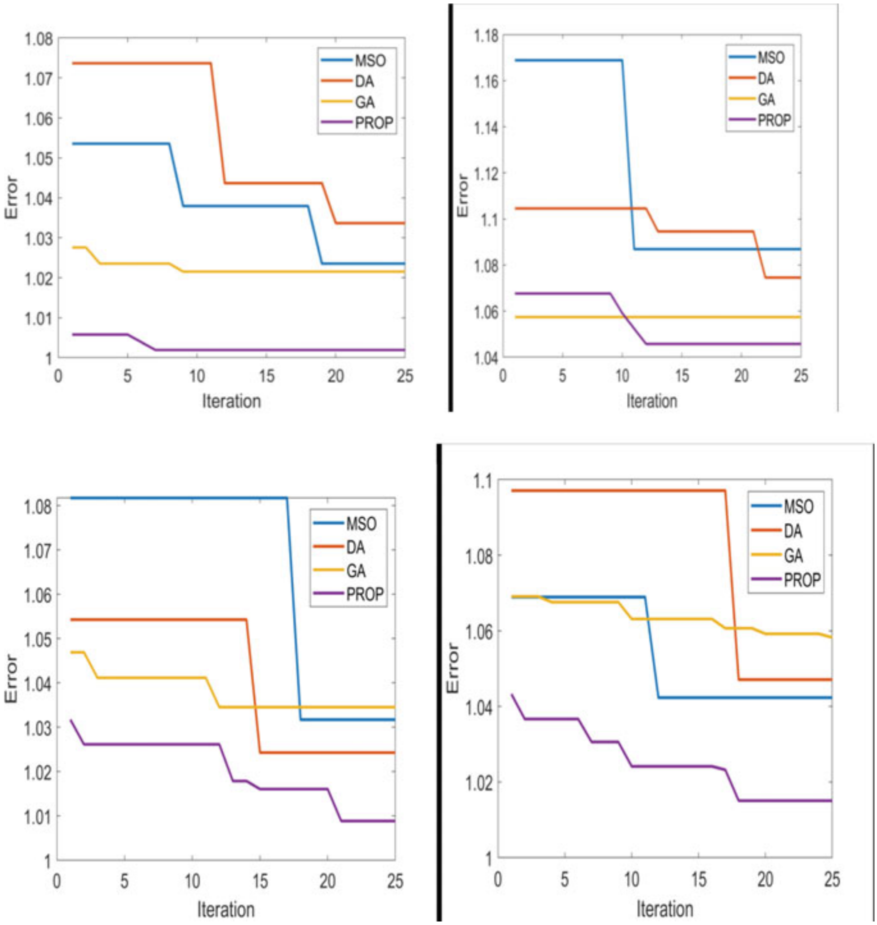


Fig. 7 (L-R) performance parameters (convergence data) showing—accuracy, precision, sensitivity, specificity

Table 3 Performance classifier sensitivity achieved with various techniques (existing and proposed)

Technique dataset	DCNN	SVM_GA	MSO	DA	DA_MSO(proposed method)
Data746	78.571	76.531	80.612	89.796	92.857
Data1625	95.137	76.531	93.658	95.137	95.983
DataShilling	95.982	76.531	94.834	95.637	96.556
DataImpens	93.878	76.531	94.286	91.837	96.735

Table 4 Performance classifier precision achieved with various techniques (existing and proposed)

Technique dataset	DCNN	SVM_GA	MSO	DA	DA_MSO (proposed method)
Data746	88.889	85.714	96.032	91.27	92.063
Data1625	42.857	85.714	42.857	50	50
DataShilling	69.369	85.714	80.18	74.775	72.973
DataImpens	56.41	85.714	66.667	64.103	56.41

Table 5 Performance classifier specificity achieved with various techniques (existing and proposed)

Technique dataset	DCNN	SVM_GA	MSO	DA	DA_MSO (proposed method)
Data746	84.615	80.645	94.048	88.889	90.099
Data1625	98.253	80.645	98.226	98.468	98.482
DataShilling	96.092	80.645	97.406	96.748	96.556
DataImpens	93.117	80.645	94.672	94.142	96.556

$$\text{Sensitivity} = {}^{T_\text{p}}\!/_{T_\text{p}} + T_\text{n}$$

3. Precision: Precision is defined as the fraction of relative instances among the retrieved instances. It is defined as the number of true positives divided by the number of true positive plus the number of false positives. The average precision achieved in the proposed methodology is 67.8%. The formula for sensitivity is given below

$$\text{Precision} = {}^{T_\text{p}}\!/_{T_\text{p}} + F_\text{n}$$

4. Specificity: It specifies the likelihood of false negatives. It is also defined as a True Negative rate. The average specificity achieved in the proposed methodology is 95.42%. Sensitivity is given by the formula

$$\text{Specificity} = {}^{T_{\text{n}}}\!/_{T_{\text{n}}} + F_{\text{p}}$$

5 Conclusion

Deep convolution neural network hybridized with moth search and dragonfly is one of the efficient tools in machine learning and optimization. The datasets are classified and trained to achieve higher accuracy with optimization techniques. HIV-protease cleavage site prediction model proposed here will assist the researcher in predicting cleavage specificity with greater accuracy as compared to other methods and automated decision support systems are one such method in medicine. The proposed will give a boost to the researcher with user-friendly, fast, and robust methods for cleavage specificity.

6 Future Scope

In this paper, data considered is replicated at some places and is very limited, replicated data should be eliminated before processing and also fast and efficient encoding technique can be devised in the future that can achieve higher accuracy in the prediction of cleavages in proteomics.

References

1. Brik A, Wong C-H (2003) HIV-I protease: mechanism and drug discovery Org Biomol Chem 1(1):5–14
2. World Health Organization. http://www.who.int/gho/hiv/en/
3. De Clercq E (2009) The history of anti retro viral: key discoveries over the past 25 years. Med Virol 19(5):287–299
4. Singh et al (2019) Clean: evolutionary based ensemble framework for realizing transfer learning in HIV-1 Protease cleavage sites prediction. Appl Intell 49:1260–1282
5. Fathi et al (2018) A genetic programming method for feature mapping to improve prediction of HIV-1 protease cleavage site. Appl soft Comput 72:56–64
6. Singh et al (2018) Evolutionary based optimal ensemble classifiers for HIV-1 protease cleavage sites prediction. Expert Syst Appl 109:86–99. https://doi.org/10.1016/j.eswa.2018.05.003
7. Ogul et al (2009) Variable context Markov chains for HIV protease cleavage site prediction. Bio Syst 96(3):246–250. https://doi.org/10.1016/j.biosystems.2009.03.001
8. Singh O, Su ECY (2016) Prediction of HIV-1 protease cleavage site using a combination of sequence, structural, and physicochemical features. BMC Bioinform 17(Suppl 17):478. https://doi.org/10.1186/s12859-016-1337-6

9. Nanni L, Lumini A (2008) Using ensemble of classifiers for predicting HIV protease cleavage sites in proteins. Amino Acids 36(3):409–416. https://doi.org/10.1007/s00726-008-0076-z(2008)
10. Jiangning Song Hao Tan,Andrew J. Perry,Tatsuya Akutsu,Geoffrey I. Webb,James C. hisstock, Robert N. Pike.: PROSPER: An Integrated Feature-Based Tool for Predicting Protease Substrate Cleavage Sites. Briefings Bioinform 20:638–658
11. Gai-Ge Wang (2016) Solar Moth search algorithm: a bio-inspired metaheuristic algorithm for global optimization problems. Memetic Comput 10:151–164
12. Li Z, Zhou Y, Zhang S, Song J (2016) Lévy-flight moth-flame algorithm for function optimization and engineering design problems. Math Probl Eng https://doi.org/10.1155/2016/1423930
13. Mirjalili S (2016) Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. Neural Comput Appl 27:1053–1073. https://doi.org/10.1007/s00521-015-1920-1
14. Mafarja MM, Eleyan D, Jaber J, Hammouri A, Mirjalili S (2017) Binary dragonfly algorithm for feature selection. In: 2017 international conference on new trends in computing sciences (ICTCS). https://doi.org/10.1109/ictcs42043
15. Rahamn S (2019) Dragonfly algorithm and its applications. Appl Sci Surv

Artificial Neural Network for Identification and Classification of Natural Body Marks

Dayanand G. Savakar, Danesh Telsang, and Anil Kannur

Abstract Natural and artificial body marks like mole and tattoos are used to identify the victims, such as suspected, and unidentified bodies like in mass death in a plane crash and the tsunami it is a very complex situation to identify the body; in recent years, classification and identification have taken a lot of attention. This paper presents the classification and identification of natural and artificial body marks like mole and tattoo. Active contour segmentation is used to segment the image. There are 28 features extracted from each mole and tattoo image, where 18(color features), 4(texture features), 6(shape features). The artificial neural network is used to classify natural and artificial body marks, and classification accuracies obtained 88.7%. The designed algorithm works based on the features that are being extracted. Several different forms of the process exist to notify the different forms of the naturally identified body marks. The designed and proposed algorithm within this paper incorporates such kind of techniques to identify the natural and artificial body marks.

Keywords Artificial neural network · GLCM—gray-level co-occurrence matrix · BPNN—backpropagation neural network

1 Introduction

Moles are little sores in the skin. They are normally tanned and found on any portion of the body that is chest, face, hands, leg, and so on. A few moles are a lot darker, and some others are skin shaded. They might be harsh, level, raised, round, and

---

D. G. Savakar · D. Telsang (✉)
Department of Computer Science, Rani Channamma University, Belagavi, Karnataka, India
e-mail: dtelsang@gmail.com

D. G. Savakar
e-mail: dgsavakar@gmail.com

A. Kannur
Department of Computer Science & Engineering, Rajarambapu Institute of Technology, Islampur, Maharashtra, India
e-mail: anilkannur1978@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_41

oval shape. Moles are normally found in individuals who are more presented to the sun instead of less uncovered [1]. Sun consumes are not moles [2]. Identification of the several different marks over the body is an application that acts as a major area mainly in the case of the medical field. In case of mass death due to floods, plane crash, tsunami, and earth quick, it is very tough to identify the person, so the victim identification bases upon their body mark like mole (natural body mark) and tattoo (artificial body mark) are useful to identify the victim [3, 4]. Every person has its identification mark like birthmark (mole), as well as if he has an artificial body mark like a tattoo [5]. Artificial body marks resemble tattoo is a type of body modification, made by embeddings permanent ink into the dermis layer of the skin to change the shade tattoo. A tattoo may be found in any part of the body that is leg, shoulder, face, hand. Promising ongoing work has established the achievability of figuring out mole and tattoo. For about 500 years people have been using steps to identify and represent themselves using tattoo [6]. At first, the tattoo was limited to certain groups, such as motorbike riders, sailors, and criminal groups, but in the modern world, it is natural to have tattoos, as about 36% of people in the world have at least 1 tattoo. For a person, an identification tattoo is a useful tool in forensic applications.

The layout of the paper is organized as follows, Sect. 2 describes the survey report of various methods. Section 3 defines the problem statement. Section 4 explains the methodology adopted for the proposed method. Section 5 depicts the results and discussions. Finally, Sect. 6 concludes the paper with future scope of the research.

2 Literature Review

The backpropagation neural network (BPNN) is used for the ANN to deploy the process of recognition and classification on almost the same images of food grain like cumin seeds, fennel seeds, mung beans, black gram, finger millet, mustard, soybeans, and black-eyed beans. The color and texture features are selected for classification, considering color, and texture 18 and 27 features are extracted, respectively, and combined color and texture 4 and 5 features are extracted [7]. Automatic detection of melanoma skin cancer using texture analysis, gray-level co-occurrence matrix (GLCM) used for feature extraction, and used multilayer perceptron network. It is a feed-forward network and author set two types of classifier, namely automatic MLP and will be traditional MLP [8].

The different neural network models are used, namely Elman's network, cascade network, and feed-forward network, watershed segmentation is used to identify the area and equivdiameter features of diverse seed varieties like lentils, wheat, redgram, groundnut, rice, bengal gram, jowar, and metagi, here 11(includes 9 area, color, and equivdiameter) features, 18 only color and 20 features are extracted, in which 18 are colors and 2 are boundaries. Elman's network remains time-consuming in the training process, and also the cascade-forward network consumes increased time in training; but it greatly reduces the memory size and feed-forward network takes less time to train accompanied with reduced memory size when compared with Elman's and cascade-forward networks [9].

To identify melanoma skin cancer, firstly edge detection technique is used to has been segmented melanoma area, then features are extracted like asymmetry, border, color, diameter (ABCD). ANN is used to classify melanoma, and finally, the backpropagation algorithm is used to identify melanoma skin cancer area [10]. Active contour and CBIR method were implemented for segmenting tattoo; for texture analysis, they used haar wavelet decomposition, and for color representation, hue-saturation-value histograms are used at the end and they got result using their new approach glocal(glocal-local) image feature methods to test the dataset [11].

The author discussed the automatic skin cancer detection system, and they used the thresholding method for segmentation and 2D wavelet transform have used for the feature extraction, and in this proposed system, they used ANN classifiers with feed-forward multilayer network and backpropagation algorithm was used for training [12].

The significant impact caused by foreign bodies in the identification and classification of the largest food grain images like rice, jowar, wheat, groundnut, and green gram in the variety of grains, plant leaves, stones, pieces of stems, weed, soil lumps, and other types of whole grains for the sake of sorting here the author used the color and texture features which are selected for classification, considering color and texture 18 and 27 feature are extracted, respectively, and combined color and texture 4 and 5 features are extracted in this work feed-forward ANN model and backpropagation algorithm are used in the training process [2, 13].

3 Problem Statement

The identification of natural and artificial body marks remains one among the challenging tasks present in image processing. Identification of criminals is a challenging job; in recent years, one of the approaches to identify the criminals based on the natural and artificial body mark like mole and tattoo. In day to day, life tattoo is attracted to everyone, some of them have a tattoo on hand, leg, neck, etc., and everyone has at least one mole in their body. It is very difficult to identify the body where mass death occurred in the case of tsunami and plane crash in that time body marks like mole and tattoo will help to identify the person. To overcome this problem, proposed a method to identify the criminals and mass death persons, the artificial neural network has been used. The proposed method steps are shown in Fig. 1.

4 Methodology

This section presents a detailed description of image acquisition, image preprocessing, segmentation, and feature extraction; finally, an artificial neural network is

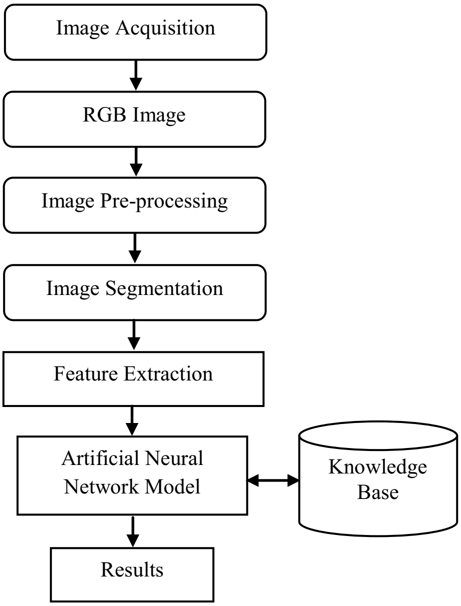


Fig. 1 Proposed system architecture

used to classify the natural and artificial body marks as explained in the subsequent section.

4.1 Image Acquisition and Preprocessing

It is a process of capturing an image from a camera, and it is a first step in the workflow sequence. The images are captured by canon 1300D camera with 1.5X to 10X lens zooming capacity and keeping object and the camera in a fixed position with maintaining clock word distance between them and some of the images were collected from an online source [14]. In the preprocessing stage, unnecessary noise is removed from the image and resized too [300 400]. The DULL RAZOR Software [15] is used to alteration, and it removes unnecessary hairs from images to improve the quality of image; it is a medical imaging software and it uses an algorithm the same as that of average filtering. The filtering technique also plays a major role in image processing to upgrade the images, sharpening, edge detection, and noise reduction. For convolution operation, spatial filtering is used as shown in Eq (1).

$$S(x, y) = \sum_{m=-M/2}^{M/2} \sum_{n=-N/2}^{N/2} h(m, n) f(s - m, y - n),$$

where $h(m, n)$ is the Gaussian filtering mask of size $M \times N$.

4.2 Segmentation (Active Contour)

Active contour is a refinement of an object boundary. It will get a curve formed by connecting their edged points, so there is no clear cut boundary for that need to connect those edge points to formed by a curved, then set of connected points, which moves to minimize a specified energy function, there is also another name for this one called as a snakes method, the contour over here is something which can flow along with the different perspectives on the image itself and then it became down to convergence, and from this particular attribute of the contour itself, it gets its name called as an active contour. The model is elected as an active model for the segmentation process. The first step, it should have an image which is extracted from its background, then, will take it as energy function like gray-level value, gradient and the initial step is segment a boundary by a general segmentation technique object with boundary has been found. The next step refines that boundary with the technique used to wiggle the snake in this it compares the pixel on each point on the boundary with energy calculated for the point in its neighborhood, it moves the boundary to neighborhood point that has the lowest energy and operates once on all points on the boundary after that repeat the iteration until it goes no for the movement the process will get stopped. The advantages of the active contour are computational efficiency and relative simplicity, and disadvantage is decision criteria worship complexity. Active contour is used in various field of segmentation based on the application. The mole and tattoo images are segmented with several segmentation techniques in that active contour has given better results as compared to the other methods [16]. Hence, in this work active contour is used for segmenting images as shown in Fig. 2.

Algorithm 1: Active contour (Region-based)-based image segmentation

```
Input:RGB image
Output:Color Segmented image
Start:
Step1: Read RGB image
Step2: Convert RGB to gray
Step3: Create a mask
Step4: Initialize number of iteration
Step6: Cover the object of an image using a region using a mask value
based on the number of iteration.
```

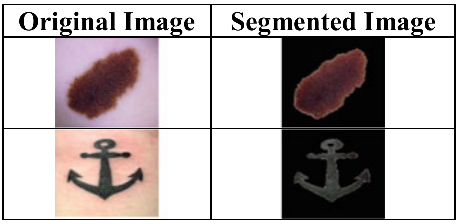


Fig. 2 Samples of the original image and segmented image using active contour segmentation technique [2]

```
Step5: Set, background = 0 and foreground = 1, (covered area of the an
image is 1)
Step7: Segment the covered area of an image
Step8: Reshape to the original image.
Stop
```

4.3 Feature Extraction

The image which is segmented by active contour is introduced to perform feature extraction (refer Table 1) in the process of identifying different body marks, which include natural mark like mole and artificial mark like a tattoo. The color features are selected because the mole and tattoo are representing different colors. The texture feature also extracted from the images based on the surface area of the image. Finally, shape features are extracted to identify the mole and tattoo images based on shape area and shape solidity, etc. From each image 28 features are extracted and stored as a feature vector. Table 1 represents the total number of features like color features (18), texture features (04), shape features (06).

The detailed description of color, texture, and shape futures are discussed in the subsequent section.

4.3.1 Color Feature Extraction

Color image consists of several components like hue, saturation, value, red, green, and blue. Three features are selected from each component like mean, variance, and range. Based on the color of the mole and tattoo image, eighteen color features are extracted from each image.

Algorithm 2: Color Features Extraction

```
Input: color image
Output: Eighteen Color features
/Start
Step 1: color image separate from the RGB components.
Step 2: Established HIS components from RGB Components
using Eqs. 2–7
Step 3: From each RGB and HSI component extracted range,
variance, and mean.
Stop.
```

$$H = \cos^{-1}\left\{\frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{1/2}}\right\} \tag{2}$$

Table 1 List of color, texture, shape, features

Sl.No	Color features
01	Red mean
02	Red variance
03	Red range
04	Green mean
05	Green variance
06	Green range
07	Blue mean
08	Blue variance
09	Blue range
10	Hue mean
11	Hue variance
12	Hue range
13	Saturation mean
14	Saturation variance
15	Saturation range
16	Value mean
17	Value variance
18	Value range
Texture feature	
19	Texture contrast
20	Texture correlation
21	Texture energy
22	Texture homogeneity
Shape feature	
23	Shape area
24	Shape convex area
25	Shape eccentricity
26	Shape solidity
27	Shape filled area
28	Shape equivdiameter

$$S = 1 - \frac{3}{(R + G + B)}[\mathrm{Min}(R, G, B)] \tag{3}$$

$$I = \frac{1}{3}(R + G + B) \tag{4}$$

$$\text{Mean} \quad \mu = \sum_{x,y} xp(x, y) \tag{5}$$

$$\text{Variance} = \sum_{x,y} (x - \mu)^2 p(x, y) \tag{6}$$

$$\text{Range} = \text{Max}(p(x, y)) - \min(p(x, y)) \tag{7}$$

4.3.2 Texture Feature Extraction

It refers to characteristics appearance of an object given by the density, shape, size, arrangement; the proportion of its elementary parts of natural and artificial body marks in a preliminary stage. The analysis of the mole and tattoo image thought the texture analysis, then extracted four features from each mole and tattoo image like contrast, correlation, energy, homogeneity.

Algorithm 3: Extraction of textural feature

```
Input: Original Image RGB
Output: 4 Texture Features
Start
Step1: 24-bit input color image separate from RGB components,
obtain the Gray-level co-occurrence matrices (GLCM).
Step 2: Compute the Co-occurrence Matrix
Step 3: GLCM features (Texture features) Energy,
Contrast, Homogeneity, correlation
Stop.
```

$$\text{Contrast} = \sum |x - y|^2 p(x, y) \tag{8}$$

Contrast refers to the degree of intensity of the pixels and its neighbor, which is calculated by using the difference in the color & brightness of the object and objects present within the similar field view is given by Eq. 8.

$$\text{Correlation} = \sum_{x,y}^{\sigma_x \sigma_y} [(xy)P(x, y)] - \mu_x \mu_y \tag{9}$$

μ_x and μ_y: mean; and σ_x and σ_y: standard deviation.

Correlation is referred as the degree of similarity between the data that relate the processing of spatial domain with the frequency-domain as given in Eq. 9.

$$\text{Energy} = \sum_{x,y} p^2(x, y) \tag{10}$$

Here, energy is referred as the degree of pixel pair repetitions at the extent, where it also leverages the image uniformity as represented by Eq. 10.

$$\text{Homogeneity} = \sum_i \sum_j \frac{P(i, j)}{1 + |i - j|} \tag{11}$$

{P (d, θ) (i, j)} (i, j)	represents the probability in occurrence of gray-level pair separated by a specified distance *d* at angle θ as given by Eq. 11.

4.3.3 Extraction of Shape Feature

The mole and tattoo features are extracted by different measures like shape area, convex area, eccentricity, solidity, filled area, equivdiameter. The area of the object present in the image is measured by shape area and measured area of the particular object as consider as a feature, stored in the feature vector. The other measures also extract the feature of the mole and tattoo images based on their properties. There are six features are extracted from each image.

Algorithm 4: Extraction of shape feature

```
Input: RGB Image
Output: 6 Shape Features
Start
Step1: RGB to Gray conversation
Step 2: Detection of an object area from an Image
Step 3: Shape Features (Shape Area, Shape Convex Area, Shape
Eccentricity, Shape Solidity, Shape Filled Area, Shape
Equivdiameter)
Stop.
```

$$\text{Shape Area A}(R) = |R| = \text{N} \tag{12}$$

It is a scalar value used to find out the number of pixels in the region as given by Eq. 12.

$$\text{Shape Convex Area } C = \left\{ \sum_{i=1}^{S} \alpha_i \chi_i \middle| (\forall i : \alpha_i \geq 0) \wedge \sum_{i=1}^{S} \alpha_i = 1 \right\} \tag{13}$$

It is a scalar value that specifies the number of pixels in the convex image as given by Eq. 13.

$$\text{Shape Eccentricity } E = \sqrt{1} - {}^{b^2}\!/_{a^2} \tag{14}$$

Here, eccentricity is obtained by using the length associated with the semimajor axis a and semiminor axis b of the object present in the image and is represented by using Eq. 14.

$$\text{Shape Solidity} = \frac{\text{Area}}{\text{Convex area}} \tag{15}$$

Solidity measure is obtained as the ratio of the object area to the area associated with the convex hull of the object Eq. 15.

$$\text{EquivDiameter} = \sqrt{\frac{4 \times \text{Area}}{\pi}} \tag{16}$$

It is a scalar quantity which denotes the circle diameter with the identical area as a region and can be computed by Eq. 16.

4.4 *Artificial Neural Network*

Artificial neural network [ANN] is the most powerful network architecture in today's word. In this work, ANN used to classify the natural and artificial body marks such as mole and tattoo based on color, texture, and shape features. There are three different layers with feed-forward architecture that are used to the formation of the neural network. The numbers of input features are considered equally with the available input layers. Outcome with number of body marks remains equal to the output layers shown in Fig. 3. From each image 28 features are extracted in that color (18), shape (6), texture (4), based on that mole and tattoo samples are recognized. The output layers have 2 nodes in all cases. The output is represented as a pattern vector of 2 bits Q(q1,q2) is set to 1 and remaining bits to 0 s, the image sample belongs to the ith type of body images. The vector Q1 (1 0), Q2(0 1) represents a mole and tattoo, respectively. The number of hidden layers is calculated using the formula as shown in Eq. (17).

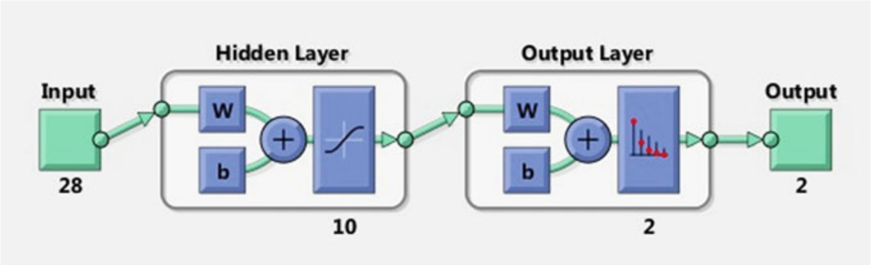


Fig. 3 Structure of ANN

$$n = \frac{I + O}{2} + y^{0.5} \tag{17}$$

where

n	number of nodes in the hidden layer
I	number of input features
O	number of outputs
y	number of input patterns in the training set.

4.4.1 Training and Testing

The identification and classification of natural and artificial body marks on different datasets like mole and tattoo are summarized as follows: Training and testing of the neural network are accomplished using body mark datasets; these datasets are divided into two parts: first one is training and testing is the second one. Totally, 266 sample images were taken for an experiment; in that, 135 images are mole and 131 images are a tattoo; 50% dataset are taken for training and 25% are taken for both validation and testing purposes. The ANN models performance for training and testing of the selected dataset shown in Fig. 4, the blue color indicates the training, red color indicates the testing, and green color indicates the validation of the experiment. The graph shows the performance of the ANN model achieving the target with many iterations in a particular time period.

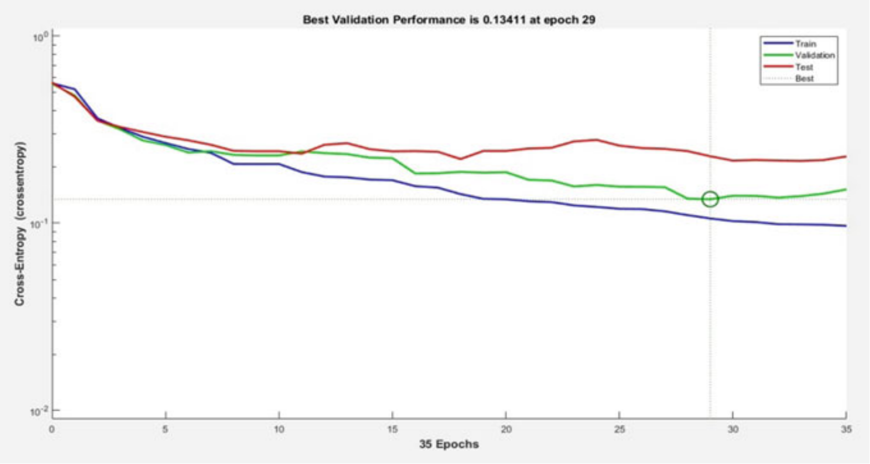
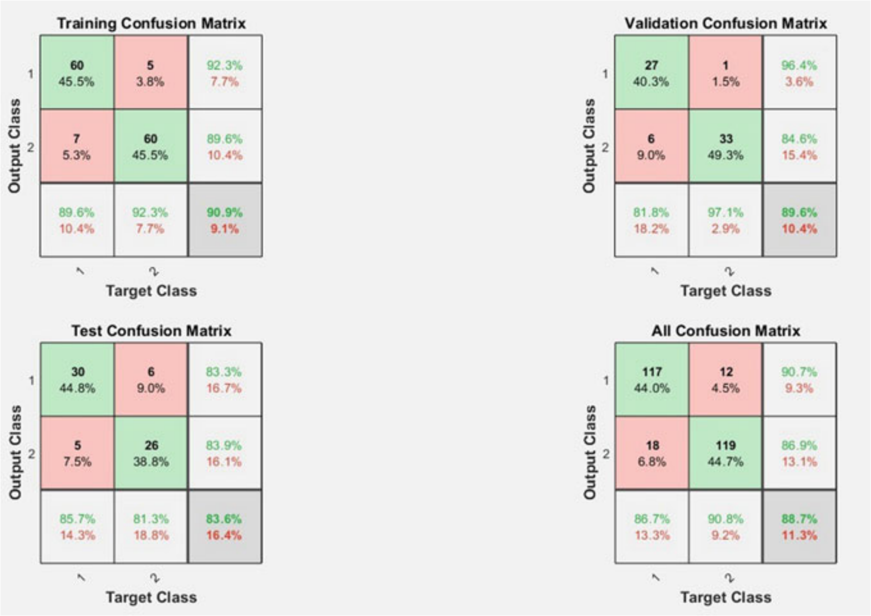


Fig. 4 Best validation performance of ANN

5 Results and Discussions

The experiment is evaluated on 266 images in which 135 contain mole images and 131 tattoo images. The total number of features are extracted from each image which is 28. The numbers of input layers are 28, the number of hidden layers is 10, and 2 output layers are set to experiment. Then 50% of the dataset is used for training and the remaining 25% for validation and 25% for testing. The proposed experimental results are shown in the confusion matrix, the prediction value of mole 86.7% true and failure case 13.3% out of 135 samples, whereas in tattoo 90.8% true and the failure case 9.2% out of 131 samples and the overall accuracy of the proposed method is 88.7% true and 11.3% failure out of 266 samples present in Fig. 5. The detailed description is given in the subsequent section.

This is a commonly used graph that summarizes the performance of a classifier's overall possible thresholds. It is generated by plotting the true positive rate (y-axis) against the false positive rate (x-axis) as you vary the threshold for assigning observations to a given class. The blue color represents class one (mole) and tan color represents the class two (Tattoo). The curves are presented in all ROC the performance of the proposed method shown in Fig. 6.



Training Confusion Matrix

Output Class \ Target Class	1	2	
1	60 45.5%	5 3.8%	92.3% 7.7%
2	7 5.3%	60 45.5%	89.6% 10.4%
	89.6% 10.4%	92.3% 7.7%	90.9% 9.1%

Validation Confusion Matrix

Output Class \ Target Class	1	2	
1	27 40.3%	1 1.5%	96.4% 3.6%
2	6 9.0%	33 49.3%	84.6% 15.4%
	81.8% 18.2%	97.1% 2.9%	89.6% 10.4%

Test Confusion Matrix

Output Class \ Target Class	1	2	
1	30 44.8%	6 9.0%	83.3% 16.7%
2	5 7.5%	26 38.8%	83.9% 16.1%
	85.7% 14.3%	81.3% 18.8%	83.6% 16.4%

All Confusion Matrix

Output Class \ Target Class	1	2	
1	117 44.0%	12 4.5%	90.7% 9.3%
2	18 6.8%	119 44.7%	86.9% 13.1%
	86.7% 13.3%	90.8% 9.2%	88.7% 11.3%

Fig. 5 Target class confusion matrix between class1 mole and class2 tattoo

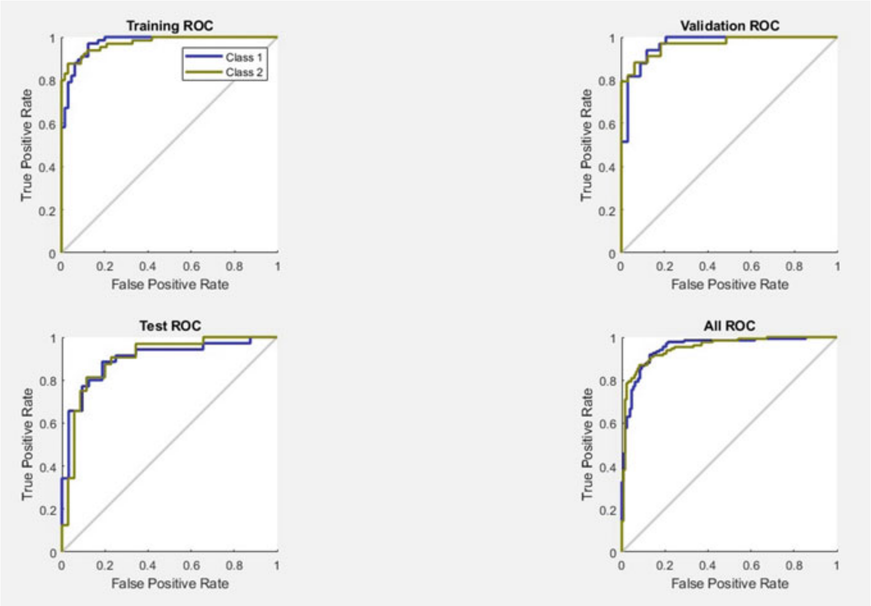


Fig. 6 Classifier accuracies for the selection of different mole and tattoo classes

5.1 Experimental Results of the Work

Mole and tattoo images are collected from various hospitals and websites; the total number of images that are considered for the experiment is 266. These images were preprocessed with the following techniques like hair removal using DULL RAZOR software and spatial filtering is used to reduce edge detection, noise reduction, and segmentation processes and finally get segmented images from the preprocessing technique. ANN toolbox is used from MATLAB platform, in that random for data division and scaled conjugate gradient for training and cross-entropy for performance, and the numbers of epochs are set to 35 iterations has been set. For the identification of mole is set to network 1 and for the tattoo is network 0. The result of mole identification is about 86.7% and the failure case is about 13.3% of 135 samples. In tattoo, the result 90.8% and the failure case 9.2% of 131 images, and the combined images overall accuracy result are 88.7%, and the failure case 11.3% of total 266 samples.

5.2 Comparative Analysis

The experiment is evaluated on artificial and natural body marks; it contains 131 and 135 images. The proposed method results are compared with other methods, whereas

Table 2 Comparison of the proposed method results with other methods

Sl.No	Methods	Results (%)
1	Neural network with backpropagation	75
2	Neural network with backpropagation	75.6
3	Auto-associative neural network	80.8
4	Auto-associative neural network	82.6
5	**Proposed method**	**88.7**

only compared with methods not with an object because the proposed method is novel. Till now no one has come across this work. The proposed method compared with the neural network with backpropagation [14, 17], an auto-associative neural network [18, 19]. The results of the proposed method are compared based on the method accuracy shown in Table 2.

The proposed method has given better results as compared to the other methods presents in the literature review.

6 Conclusion and Future Scope

The proposed method has shown that the classification and identification of body marks using ANN. The proposed method obtained 86.7% accuracy for mole and 90.8% for tattoo and overall accuracy is about 88.7% when both images were combined. Using color, texture, and shape features set for mole and tattoo, respectively. The proposed method is very useful to classification and identification between natural and artificial body marks like mole and tattoo. The proposed research work is more relevant to the body marks type recognition and classification, and it will integrate both pattern recognition and image processing technologies. The future research work will be progressed to classify and identify the artificial body marks and accidental body marks, which is useful in several areas like identifying the criminal and identify the bodies where mass death occurred.

References

1. Savakar DG, Telsang D (2016) A survey on identification and analysis of body marks. Int J Innov Res Sci Eng Technol 5(5) (May)
2. A Brief History of Tattoos. http://www.designboom.com/history/tattoo_history.html
3. Jain AK, Lee JE, Jin R (2007) Tattoo-ID: automatic tattoo image retrieval for suspect and victim identification. PCM, pp 256–265
4. Lee J-E, Jain A, Jin R (2008) Scars, Marks, and Tattoos (SMT): soft biometric for suspect and victim identification. In: Biometrics symposium, Sept 2008
5. Ngan M, Grother P (2015) Tattoo recognition technology—challenge (Tatt-C): an open tattoo database for developing tattoo recognition research. National Institute of Standards and Technology (NIST)

6. Lee T, Ng V, Gallagher R, Coldman A, McLean D (1997) Dullrazor A software approach to hair removal from images. Comput Biol Med 27(6):533–543
7. Savakar DG (2012) Recognition and classification of similar looking food grain images using artificial neural networks. J Appl Comput Sci Math 13(6) (Suceava)
8. Sheha MA, Mabrouk MS, Sharawy A (2012) Automatic detection of melanoma skin cancer using texture analysis. Int J Comput Appl 42(20) (March)
9. Kannur A, Kannur A, Rajpurohit VS (2011) Classification and grading of bulk seeds using artificial neural network. Int J Mach Intell (IJMI) 3(2):62–73
10. Kanimozhi T, Murthi A (2016) Computer-aided melanoma skin cancer detection using artificial neural network classifier. Singaporean J Sci Res (SJSR) J Sel Areas Microelectron (JSAM) 8(2):35–42
11. Acton ST, Rossi A (2008) Matching and retrieval of tattoo images: active contour cbir and glocal image features. In: IEEE SSIAI, Mar 2008
12. Abdul Jaleel J, Salim S, Aswin RB (2012) Artificial neural network-based detection of skin cancer. Int J Adv Res Electric Electron Instrum Eng 1(3) (Sept)
13. Anami BS, Savakar DG (2009) Effect of foreign bodies on recognition and classification of bulk food grains image samples. J Appl Comput Sci 6(3) (Suceava)
14. Mhaske HR, Phalke DA (2013) Melanoma skin cancer detection and classification based on supervised and unsupervised learning. Circuits, Controls and Communications (CCUBE). In: 2013 Int. Conf. IEEE Dec 2013, pp 1–5
15. Tschandl P et al (2018) The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci Data 5:180161. https://doi.org/10.1038/sdata.2018.161
16. Savakar DG, Telsang D, Kannur A (2020) Comparative analysis of segmentation technique for different body marks. Int J Adv Sci Technol (IJAST) 29(4):8605–8614
17. Mahmoud MKA, Al-Jumaily A, Takruri M (2011) The automatic identification of melanoma by wavelet and curvelet analysis: study based on neural network classification. In: Hybrid intelligent systems (HIS), 2011 11th Int. Conf. IEEE, Dec 2011, pp 680–685
18. Lau HT, Al-Jumaily A (2009) Automatically early detection of skin cancer: study based on neural network classification. In: International conference of soft computing and pattern recognition, IEEE, pp 375–380
19. Srivastava S, Sharma D (2016) Automatically detection of skin cancer by classification of neural network. Int J Eng Tech Res 4(1):15–18

Improvised Distributed Data Streaming Scheduler in Storm

J. Geetha, D. S. Jayalakshmi, Riya R. Ganiga, Shaguftha Zuveria Kottur, and Tallapalli Surabhi

Abstract Apache Storm is one of the most widely used platforms for processing of data streams due to its properties of being distributed, highly scalable, and fault-tolerant. It provides real-time processing, is fast and stateless, and uses master–slave architecture with ZooKeeper. In the Hadoop ecosystem, Apache Storm is the one that fills the present real-time functionality and provides strong coupling with many tools and technologies. In Storm framework, the Storm default scheduler is commonly used to schedule the task or data to be processed, whose basis for scheduling the task is the time quanta or time slots, which leads to increase in context switches and longer response time. In Storm default scheduler, the workload of a topology equally distributed among worker processes or Java virtual machine (JVM) all over the cluster using a simple round-robin algorithm without considering any priority or criteria. The proposed algorithm addresses the above-specified issues. An improvised the custom Storm scheduler was developed where the scheduling is based on the workload, which is calculated based on the total memory utilized per task and the total processing unit utilized per task, thereby resulting in lesser context switches and faster response time of distributed streaming applications.

Keywords ZooKeeper · Java virtual machine (JVM) · Self-timed periodic (STP) model

1 Introduction

Apache Storm is a distributed stream processing computation framework. It utilizes spouts and bolts to specify information sources and manipulations to allow distributed batch processing of streaming data. The topology consists of spouts and bolts which

---

J. Geetha (✉) · D. S. Jayalakshmi · R. R. Ganiga · S. Z. Kottur · T. Surabhi
Department of Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore 560054, India
e-mail: geetha@msrit.edu

D. S. Jayalakshmi
e-mail: jayalakshmids@msrit.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_42

are interconnected. It is a directed graph consisting of vertices which are computational nodes and edges which are data streams. A simple topology begins with spouts, emitting the data to one or more bolts. In the topology, bolt replicates a node having the small processing logic unit and the output of a bolt can be emitted as an input into another bolt. Storm topology stops only after it is killed. The data that is generated continuously by different sources is called Streaming data. Such data has to be processed by an incremental procedure using Stream processing techniques without allowing access to all of the data. In addition, data drift must be considered, which means that the properties of the stream may change over time. The schedulers existing in the Storm are: Isolation scheduler, Multitenant Scheduler, Default Scheduler (also known as the Even Scheduler), and Pluggable Scheduler. The isolation scheduler provides an easy and safe mechanism for sharing Storm cluster resources among many topologies. Within the Storm cluster, the isolation scheduler helps to allocate or reserve the dedicated sets of nodes for topologies. A Storm default scheduler assigns component executors as equally as possible between all the workers. In the pluggable scheduler, one can replace the default scheduler by implementing their scheduler which assigns executors to workers.

2 Background

Apache Storm [1] is a system used for real time, distributed big data processing [2, 3]. It is being used at for analysis of social media at Twitter [4] and Klout, for weather data [5], in telecommunication industries [6], etc. A Storm cluster consists of master–slave architecture where the Nimbus daemon is run on the master node and Supervisor daemon is run on the worker nodes. Coordination between the Supervisors and Nimbus daemons is managed by the ZooKeeper cluster as shown in Fig. 1. The responsibility of the Nimbus daemon includes distribution of code around the cluster, task assignment to machines, and failure monitoring. The supervisor daemon listens for work assigned to its worker machine and starts and stops worker processes as necessary based on what Nimbus has assigned to it. To perform real-time computations, topologies are created which is a graph of computation. Each worker process executes a part of the topology. The node of a topology possesses processing lodges and the edges indicate the direction of data transfer between nodes [7].

A stream is an unbounded sequence of tuples, which is the data that the topology works upon. Storm provides the primitives for the transformation of streams in a reliable and distributed way. The basic primitives provided by Storm are “spouts” and “bolts.” A spout is a source of streams and a bolt consumes input streams, performs some processing, and possibly emits new streams. An executor is a thread created by a worker process.

For a specific spout or bolt, the executor can run one or more tasks. Networks of spouts and bolts form a topology. Each node is a spout or bolt. Edges indicate the subscription of streams by the nodes which are either spouts or bolts. When a tuple

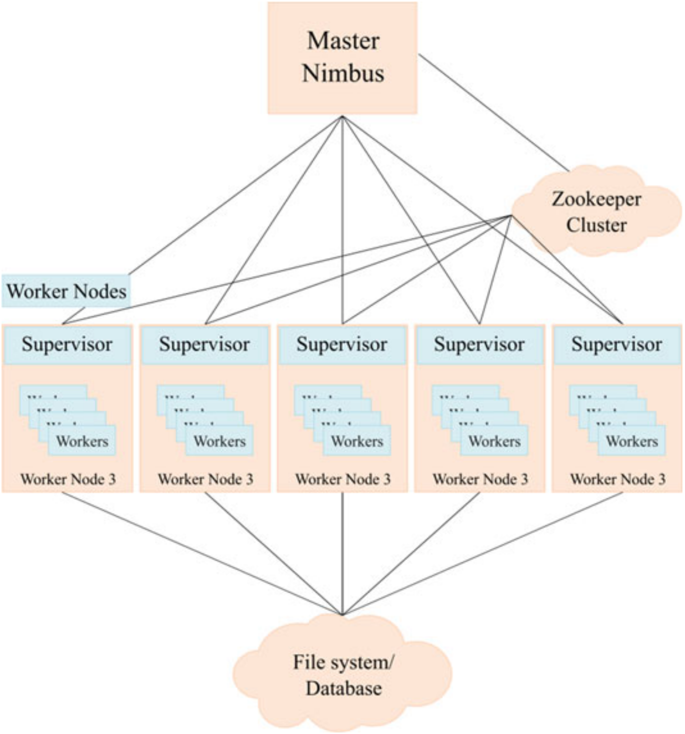


Fig. 1 Apache Storm Cluster Architecture depicting master node with Nimbus daemon, worker nodes with supervisor daemon and various worker slots; coordination is managed by ZooKeeper cluster

is generated, it is forwarded to all the nodes which have subscribed to the particular stream. A topology runs forever until it is killed.

The schedulers existing in the Storm are: Isolation scheduler, Multitenant Scheduler, Default Scheduler (also known as the Even Scheduler), and Pluggable Scheduler. The Storm default scheduler aims to produce an even allocation and hence utilizes the round-robin strategy. It iterates through the executors in the topology and uses round-robin to assign them to the configured number of workers in the topology. Then, according to slot availability of the nodes, these workers are then assigned to worker nodes. Storm provides the capability to implement custom schedulers which need to implement the "prepare" and "schedule" methods of the IScheduler interface. The custom scheduler takes as input the topology structure as a weighted graph and a set of additional user-defined parameters. Then a deployment plan or schedule is computed which outlines the mapping of executors to workers and the assignment of workers to slots, as shown in Fig. 2. The Storm scheduler can either be made to execute periodically or only when the topology is submitted [8].

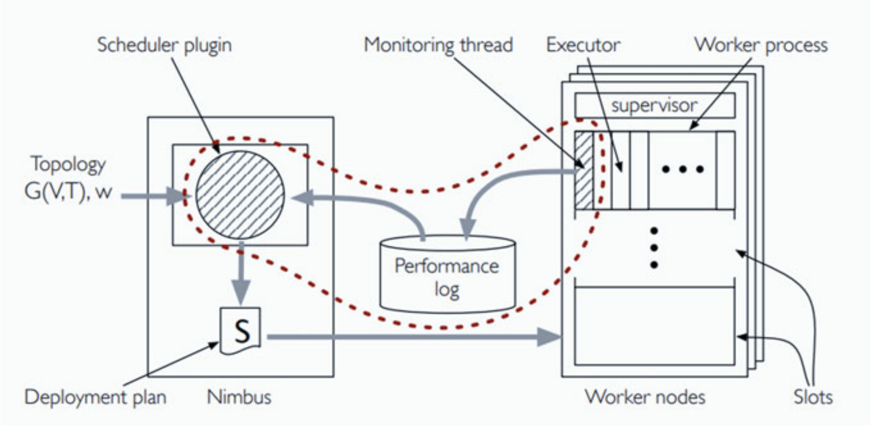


Fig. 2 Apache Storm Architecture with a depiction of scheduling methodology—the scheduler plugin holds the scheduling logic, based on the topology a deployment plan (schedule) is created which is then executed by the communication between Nimbus and Supervisors [8]

3 Related Work

The problems with the current scheduling practices in Apache Storm and associated traffic-aware online scheduling challenges [9] are discussed in this paper. The conclusions inferred from the experiments include the significant impact of internode/interprocess traffic on performance and the counter-effect of overloading a worker node. Along with the default scheduler, two more schedulers are used: an offline scheduler and an online scheduler. The design and implementation of T-Storm for enabling traffic-aware online scheduling were done by revision of certain Storm components and the addition of a schedule generator, a custom scheduler and load monitors. Scheduling in T-Storm operates as follows: (1) Periodic collection of runtime workload and traffic load information is done by the load monitor and stored in the database. (2) The schedule generator periodically reads load information from the database and computes a schedule using a traffic-aware online scheduling algorithm. (3) The current scheduler is regularly fetched by the custom scheduler and executed by assigning executors. The desirable properties of T-Storm are: Consolidation for Cost Reduction, Traffic-aware Online Scheduling, Worker Node, Optimization for Re-assignment Overhead, Optimization for the number of slots, Hot-Swapping of Scheduling Algorithms, and 12 Storm User Transparency. The experiments are conducted using Word Count, Throughput Test, and Log Stream Processing.

For dynamic scheduling required in streaming applications, stable scheduling algorithms are preferred over efficient scheduling algorithms. In [10], a Stable Online scheduling strategy with Makespan Guarantee (SOMG) is discussed which includes mathematical dependencies between system stability, resource utilization, and response time; specification of acceptable response time objectives and high system stability requirements to be met; structural optimization of a data stream graph by quantification and adjustment of vertices and heuristic critical path scheduling of a data stream graph. The system architecture consists of space for graph, Storm, hardware, and user and the topology includes subsystems of Nimbus, ZooKeeper, and Supervisor. The performance evaluation is done using a simulation environment consisting of the master node, ZooKeeper node and 16 worker nodes, and parameter settings based on system throughput, response time and vertices live migration ratio.

For prediction of a given scheduling solution, the accurate prediction of the average tuple processing time of an application can be done using a topology-aware [11] method. For scheduling, assigning threads to machines using an effective algorithm under the guidance of prediction results is used. To evaluate and validate, Storm was tested with three representative applications. The proposed framework consists of modules such as performance predictor, data collector, data store, schedule generator, time synchronizer, data pre-processor, and custom scheduler. Performance prediction, schedule generation, and the custom scheduling constitute for the important modules of the framework proposed. Assigning threads to machines is served by a prediction which is a ground rule for a scheduling algorithm and the objective of the scheduling problem over an application graph G is to minimize the average tuple processing time. From the experimental results, the following observations are made: The accuracy prediction for the Word Count Topology (Stream Version) is 83.3%; the accuracy of prediction for the Log 13 Stream Processing Topology of 84% is slightly lower than the accuracy of individual predictions; And for the Continuous Query Topology, the prediction accuracy is 85.2%.

Self-Timed Periodic [12] is an execution model, which is a combination of self-timed scheduling and periodic scheduling. In self-timed scheduling, actors are fired soon after the data dependency is observed. This schedule, since a long time, is noticed as the most appropriate policy for streaming applications. To resolve the problem with respect to the Periodic Schedule model of static nature regarding latency increase in unbalanced graphs can be resolved by using the self-timed periodic schedule. Self-timed periodic schedule is a hybrid execution process model based on the integration of self-timed schedule and periodic schedule while considering varying IPC times. To demonstrate the performance of self-timed periodic schedule model, two classes of self-timed periodic schedules based on two different granularities are used; the first schedule is denoted as STPIqi, and the other schedule is denoted as STPI ri. Under periodic scheduling, the effect of self-timed periodic schedule can be modified by putting back with actor period in every stage with its worst-case execution time. In worst case, execution time is the summation of all actor's computational and communication time. The performance of newly proposed scheduling policy is evaluated based on the experiment conducted on a set of 10 real-life streaming applications: DCT, FFT, Beamformer, Filterbank, MP3, Sample-rate, H.263 Encoder, H.263 Decoder, Bipartite, and Satellite.

Distributed Stream Processing Systems [13] are gathering interest as they provide the capability to scale invisibly by using the distributed resources of a cloud environment. One of the major challenges faced by DSPS is the development of effective scheduling mechanisms which can manage resources and their allocation to the concurrently running data analytics tasks. The proposed model aims to develop a scheduling mechanism that allocates resources to applications according to their priority by attaching it to the vertices of the processing flow graphs which represent the processing components and mapping operator priority to tuple priority. To achieve this, a meta-scheduler is used whose task is to intercept tuples, assign them priority according to their destinations, reorder them based on priority, and then forward them to their destination. The approach is implemented on Quasit, an open-source DSPS, by developing two new QoS specifications, corresponding to priority schemas: absolute or proportional priority and priority specifications. The experimental evaluation based on meta-scheduler overhead shows that the performance of the meta-scheduler depends mainly upon the number of priority classes. Evaluation of vehicle traffic proved that using the meta-scheduler allowed the application to handle the high input rate of the data streams.

Heuristic scheduling [14] algorithm operates by using graph partitioning algorithms and mathematical optimization software package to obtain reliable and efficient tasks involving more communication. Though optimal solution for efficient scheduling exists, for large size problems characterized by huge search space and high complexity of computation makes it infeasible to use optimization scheduling technique to obtain such an optimized solution. I-Scheduler is an iterative heuristic-based graph partitioning algorithm focusing on cutting down the total number of tasks for the specified problem so that the optimization software can operate in the scheduled period. Here, our performance metric is the average throughput (average of the total number of instances) accomplished in every bolt's task per 10 s period. The two real-world topologies used are—Smart Homes application based on Load Prediction model technique and NYC Taxi Data operated by Top Frequent Routes method. Although for the top frequent Routing topology there is no significant performance upgrade by the usage of I-Scheduler on top of R-Storm, it(R-Storm)needs a substantial amount of tuning. On comparison, it has been observed that similar performance results were achieved using I-Scheduler without requiring such extensive tuning.

Suitability of scheduling algorithms depends on the parameter of importance [15–17]. Real-time applications can generally tolerate failed messages with a higher margin compared to other applications. With the requirement of higher quality in streaming applications, this tolerance reduces and hence scheduling should also take into account this development.

4 Proposed Algorithm

The objective of this project is to develop an improvised scheduler in the Storm for distributed data streaming application based on the workload, thereby resulting in faster response time of distributed streaming applications. The following are some of the objectives of the project:

1. To optimize the existing Storm scheduler to get faster response time and less waiting time.
2. To reduce the number of context switches, by including the workload calculation criteria for the assignment of executors to worker slots.
3. To prioritize the tasks, such that a task with the maximum workload is executed first.

The algorithm is a modification to the existing Even Scheduler. Here the task having maximum workload is taken and scheduled. A task is defined as either the execution of a spout or a bolt. There exists workload calculation before choosing which task has the highest priority. The score is calculated as the maximum among Total memory consumed per task and total processing unit consumed per task. Figure 3 depicts the algorithm (Fig. 4).

The steps of the entire algorithm are explained below:

1. Score to Executor Mapping—From a cluster, topologies are selected which need scheduling, then for each such topology, the TotalMemoryPerTask and TotalCpuPerTask required are calculated, which are available in the topology specifications. Find the maximum among them; this is known as score calculation. Then the score is mapped to an executor of the topology arbitrarily. Hence, the first step of Score to executor mapping.
2. Scheduler Assignment—From the above mapping, the task with the maximum workload is chosen, and check for free WorkerSlot, which has already been

```
---------------------------------------
Algorithm 1 Proposed Algorithm
---------------------------------------
Procedure BASIC_SCHEDULER <parameters: Topologies Clus-
ter>
Loop1 - Topology wise:
  Choose the Topology which needs Scheduling from the
  cluster
  Loop2 - Executer wise:
    Calculate the score <- max (TotalMemReqTask,
    TotalCpuReqTask)
  Loop2 end
  Loop3 - WorkerSlot wise
    Map the executors to workerSlots only if:
      Executer is free && WorkerSlot is free
    End if
  Loop3 end
  Loop4: Reverse Mapping and Cluster Assignment
    Map nodePort and Executor
    Assign topology to cluster
  Loop4 ends
Loop1 ends
Scheduled Topologies are returned
```

Fig. 3 Proposed algorithm of custom scheduler

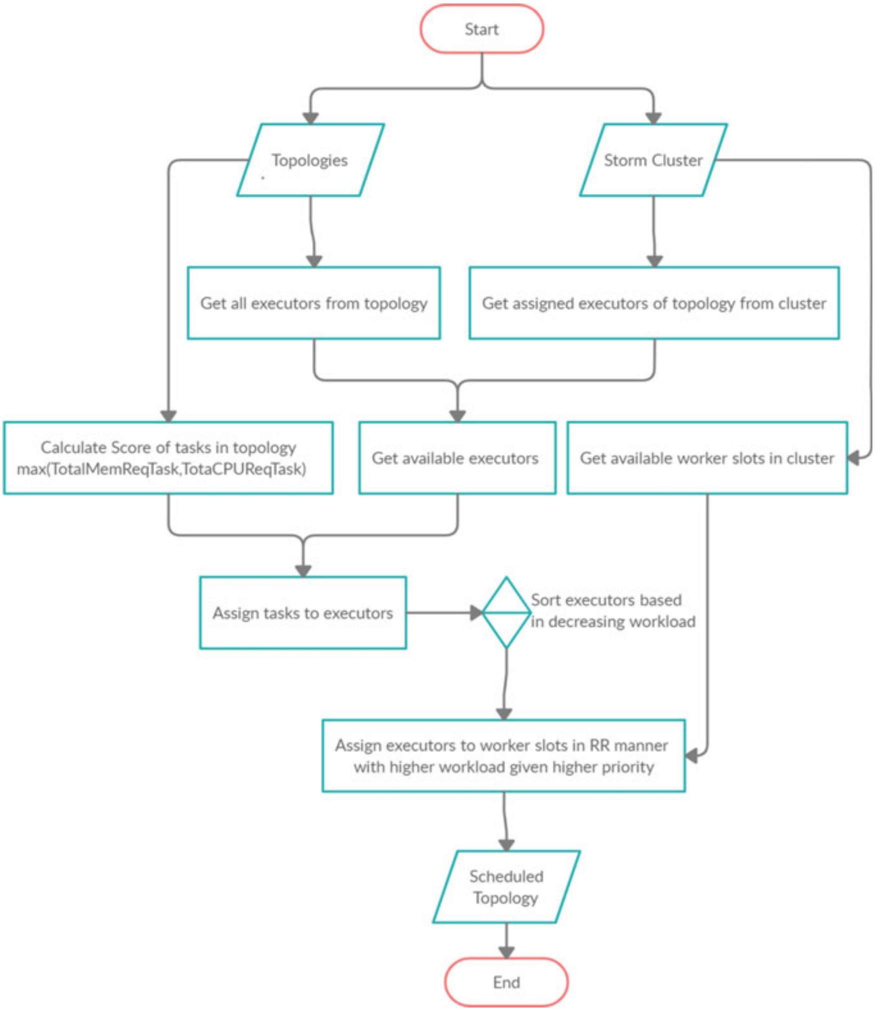

Start
Topologies
Storm Cluster
Get all executors from topology
Get assigned executors of topology from cluster
Calculate Score of tasks in topology max(TotalMemReqTask,TotaCPUReqTask)
Get available executors
Get available worker slots in cluster
Assign tasks to executors
Sort executors based in decreasing workload
Assign executors to worker slots in RR manner with higher workload given higher priority
Scheduled Topology
End

Fig. 4 Flow diagram for the proposed scheduling algorithm

mapped to an Executor. This is known as Scheduler Assignment. This is iterated till either the slots are not occupied or the tasks are finished. The task with a higher workload has higher priority.

3. NodePort to Executor Mapping—Here nodePorts (of the worker nodes) are mapped to the WorkerSlots which in turn are mapped to the Executors. This is also known as Reverse Mapping of WorkerSlots and executors.
4. Cluster Assignment of Topologies—At the end, the scheduled topologies are assigned to the cluster. This topology assignment is carried out in the Cluster.java class. There are many ways of assigning the topology back to the cluster. After this assignment, the topology is said to be scheduled.

The proposed custom Storm Scheduler involves a calculation of the workload based on memory and CPU utilized per task of a topology. This reduces the context switch present in round-robin, which is based on time quanta. This leads to a reduction in response time, increase in the number of acknowledgement of the stream and a decrease in the average latency.

5 Results

The *FastWordCountTopology* was executed with both the Default Scheduler and the Custom Scheduler which takes in a string input and calculates the *WordCount* in a non-traditional way. The metrics are compared across the uptime. The Metrics include Acknowledged, Average Latency, Failed, and Acknowledged per second which are obtained from the *FastWordCountTopology.printMetrics* method. Uptime represents the time duration for which the topology has been running and Acked denotes the number of messages/stream packets acknowledged in a given window. Similarly, Failed denotes the number of messages failed in a given window. The topology keeps track of all statistics related to it for a given time window and has an attribute called *completeLatency* which denotes the total latency for processing the message in a given window. The average latency is calculated by taking the average of the sum of *completeLatency* times the Acked messages in a window, for all windows for the corresponding uptime. The below snapshots are a comparison of the default scheduler and the custom scheduler based on different metrics have been chosen.

From Fig. 5, it can be inferred that the Custom Scheduler can acknowledge more and more stream packets as the uptime increases as compared to that of Default Scheduler. From Fig. 6, it is inferred that the Custom Scheduler and Default Scheduler have approximately the same Average Latency. Though it is observed that initially,

Fig. 5 Asked-uptime for default versus custom scheduler

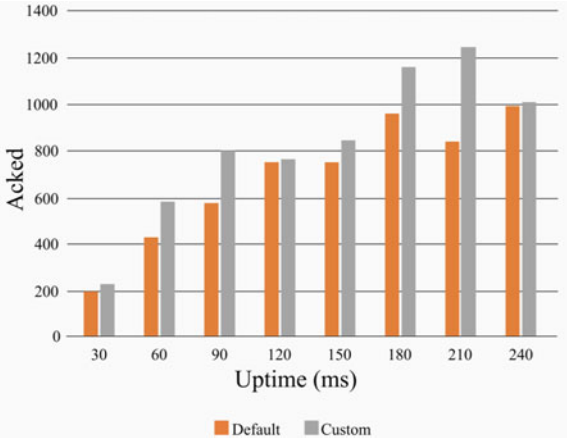


Fig. 6 Average latency of default versus custom scheduler

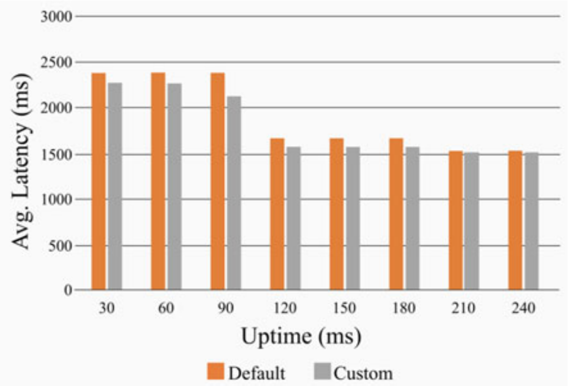
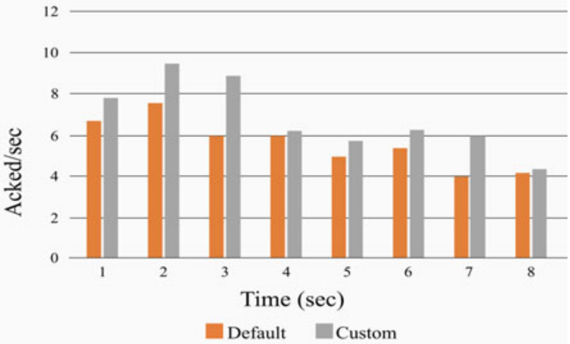


Fig. 7 Asked per second for default versus custom scheduler



the latency of the Custom Scheduler was slightly lower than the Default Scheduler; later on, it is seen to become almost the same.

In Fig. 7, the comparison of the Acknowledged per second across uptime in Storm default scheduler and Storm custom scheduler can be seen. This shows that the rate at which the custom scheduler acknowledges the stream packets increases faster than the default scheduler as uptime increases.

Acknowledgements in Custom Scheduler are decently higher that Default Scheduler, similarly average latency is slightly less in the custom scheduler. So when Ack/sec is considered, it is observed that in some cases Custom Scheduler outperforms the Default Scheduler.

6 Conclusions and Future Work

The project presents a new Custom Storm Scheduler, which is a modification of the Default Storm Scheduler. The proposed Storm scheduler works on the concept of finishing off the larger tasks of a topology first. The size of the tasks is calculated based on the workload of memory and CPU per task of a topology. This reduces the context switch present in round-robin, which is based on time quanta. The Storm custom scheduler designed as compared to that of the Storm default scheduler, it has been noted that; Storm custom scheduler made visibly outperforms the Storm default scheduler in the number of acknowledgements made per second. Storm custom scheduler still has approximately the same average latency when analyzed in comparison with the Storm default scheduler. This leads to a future scope of reducing it henceforth. The Storm custom scheduler was successful in acknowledging more streams than the Storm default scheduler, hence leading to quick data processing in steps after scheduling.

The proposed Storm custom scheduler does not differ much in average latency in comparison with the Storm default scheduler. Thus, in future, a more advanced methodology can be introduced to overcome this issue. In future, the Storm scheduler can be further modified to make it more fine-grained by introducing a new grouping methodology or any other useful criteria based on topology or source generating data just like workload calculation in the current project.

References

1. Iqbal MH, Soomro TR (2015) Big data analysis: Apache storm perspective. Int J Comput Trends Technol 19(1):9–14
2. Van Der Veen JS, van der Waaij B, Lazovik E, Wijbrandi W, Meijer RJ (2015) Dynamically scaling apache storm for the analysis of streaming data. In: 2015 IEEE first international conference on big data computing service and applications. IEEE, pp 154–161 (March)
3. Yang W, Liu X, Zhang L, Yang LT (2013) Big data real-time processing based on storm. In: 2013 12th IEEE international conference on trust, security and privacy in computing and communications. IEEE, pp 1784–1787 (July)
4. Toshniwal A, Taneja S, Shukla A, Ramasamy K, Patel JM, Kulkarni S, Jackson J, Gade K, Fu M, Donham J, Bhagat N (2014) Storm@ twitter. In: Proceedings of the 2014 ACM SIGMOD international conference on Management of data, pp 147–156 (June)
5. Gu S, Yao L, Tunc C, Akoglu A, Hariri S, Ritchie E (2016) An autonomic workflow performance manager for weather research and forecast workflows. In: 2016 international conference on cloud and autonomic computing (ICCAC). IEEE, pp 111–114 (September)
6. Schaefer C, Manoj PM (2015) Enabling privacy mechanisms in apache storm. In: 2015 IEEE international congress on big data. IEEE, pp 102–109 (June)
7. Apache Software Foundation (2019) Tutorial Apache Storm Version 2.2.0, Apache Software Foundation 2019, viewed on Sept 2020. https://storm.apache.org/releases/current/Tutorial.html
8. Aniello L, Baldoni R, Querzoni L (2013) Adaptive online scheduling in storm. In: Proceedings of the 7th ACM international conference on Distributed event-based systems, pp 207–218 (June)
9. Xu J, Chen Z, Tang J, Su S (2014) T-storm: traffic-aware online scheduling in storm. In: 2014 IEEE 34th international conference on distributed computing systems. IEEE, pp 535–544 (June)
10. Sun D, Huang R (2016) A stable online scheduling strategy for real-time stream computing over fluctuating big data streams. IEEE Access 4:8593–8607
11. Li T, Tang J, Xu J (2016) Performance modeling and predictive scheduling for distributed stream data processing. IEEE Trans Big Data 2(4):353–364

12. Dkhil A, Do XK, Louise S, Rochange C (2015) A hybrid scheduling algorithm based on self-timed and periodic scheduling for embedded streaming applications. In: 2015 23rd Euromicro international conference on parallel, distributed, and network-based processing. IEEE, pp 711–715 (March)
13. Bellavista P, Corradi A, Reale A, Ticca N (2014) Priority-based resource scheduling in distributed stream processing systems for big data applications. In: 2014 IEEE/ACM 7th international conference on utility and cloud computing. IEEE, pp 363–370 (December)
14. Eskandari L, Mair J, Huang Z, Eyers D (2018) Iterative scheduling for distributed stream processing systems. In: Proceedings of the 12th ACM international conference on distributed and event-based systems, pp 234–237 (June)
15. Eskandari L, Huang Z, Eyers D (2016) P-Scheduler: adaptive hierarchical scheduling in apache storm. In: Proceedings of the Australasian computer science week multi conference, pp 1–10 (February)
16. Liu Xm Buyya R (2017) D-storm: dynamic resource-efficient scheduling of stream processing applications. In: 2017 IEEE 23rd international conference on parallel and distributed systems (ICPADS). IEEE, pp 485–492 (December)
17. Peng B, Hosseini M, Hong Z, Farivar R, Campbell R (2015) R-storm: resource-aware scheduling in storm. In: Proceedings of the 16th annual middleware conference, pp 149–161 (November)

Efficient Mining of Rare Itemsets

Shwetha Rai, Geetha M., and Preetham Kumar

Abstract Itemset mining discovers interesting patterns in the dataset. The itemset may be frequent or it can be rare based on its occurrence in the database. It has been observed that most of the algorithms are designed for mining frequent itemsets. However, discovery of rare itemsets is equally important since they play a major role in making decisions in some situations. The efficiency of the algorithms depend on the way in which the data structures are designed to store and retrieve the data. Hyperlinked Rare Pattern Mining algorithm discovers all rare itemsets and is suitable for sparse dataset. In this algorithm item_id and its support count are stored in Support and Header tables. This redundancy is removed in the proposed algorithm to improve the time efficiency. An experimental analysis is conducted to discover the rare itemsets. It is observed that while there is an improvement in time efficiency, there is a tradeoff for space efficiency.

Keywords Association rule mining · Data engineering · Data structure · Hyper-link · Rare itemsets · Redundant data · Sparse data

1 Introduction

Data mining, an important field in computer science, is the process of discovering patterns, that are both interesting and useful, and relationships among the patterns in huge amount of data [5]. Pattern mining can be used to extract patterns that are

---

S. Rai (✉) · G. M.
Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka 576104, India
e-mail: shwetha.rai@manipal.edu

G. M.
e-mail: geetha.maiya@manipal.edu

P. Kumar
Department of Information and Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education,Manipal, Karnataka 576104, India
e-mail: preetham.kumar@manipal.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_43

either frequent or rare from the bulky dataset. Rare Pattern mining is a technique used to extract rare patterns, the patterns of which the support count is between two user-defined thresholds, freqSup and rareSup, from the dataset.

Most of the research works on pattern mining are carried out only on the extraction of frequent patterns for a considerable period disregarding the extraction and analysis of rare patterns. In recent years, pattern mining techniques showed a key role in solving numerous tasks in Data mining. Substantial extent of research has been accomplished for the discovery of rare itemsets.

Data structures in the implementation of pattern mining algorithms play a key role in discovering the itemsets of interest. A tree data structure is more suitable when the dataset is dense but the same data structure may not be efficient to store itemsets from a sparse dataset. Hyper-Linked RPM algorithm is an attempt to store sparse dataset in the main memory.

2 Literature Survey

Since the inception of the research work on rare pattern mining, a significant extent of research has been accomplished for the discovery of rare patterns. Different authors have proposed different techniques to extract patterns from the dataset. There are different techniques for discovery of rare patterns such as level-wise approach or candidate generation approach like Apriori [2] or a non-candidate generation method, like FP-Growth [7] which is efficient than candidate generation algorithm. The rare itemsets are also discovered based on other constraints such as multiple support counts [3, 8–12] and based on high utility, a weight given to each item in the database [6, 15].

An Apriori approach, a candidate generation method, for the discovery of rare itemsets was first proposed in the paper [13]. Algorithms such as ARIMA [16] and AfRIM [1] used a single support value for identifying and extracting the rare itemsets. It is known from the existing literature for rare itemsets mining that there are many drawbacks of in this approach and hence to overcome those drawbacks other techniques were developed.

Rare Pattern Tree or RP-Tree, implemented in [17] is a tree based algorithm to store rare itemsets. This algorithm supports two threshold values and the transactions that has at least one rare itemset were considered for further processing. RP-Tree algorithm showed a better performance when multiple support thresholds were used for the itemsets [3]. Since dense datasets contain many frequent itemsets, tree-based approaches are considered to be more suitable to discover frequent itemsets. The tree-based approaches that gave a good result for the data having long patterns, failed for data having short patterns.

In the paper [4], a queue-based approach to mine infrequent patterns, “Hyper-Linked Rare Pattern Mining (HLRPM)”, was proposed. The authors claimed that it generates better results when the dataset provided is sparse and has long patterns. It is space efficient when compared to other tree based rare pattern mining algorithms. The algorithm uses hyper-link data structure [14] to store the rare transactions. Since it employs memory-based data structure, it outperforms the level-wise and pattern growth approaches in various cases.

Each of the techniques discussed in the previous section has its own advantages and drawbacks. In the paper [4], as claimed by the author, the algorithm is faster compared to other approaches but it takes two database scans and one additional scan through the support count table which affects the time consumption.

3 Research Methodology

3.1 HLRPM Algorithm

Algorithm 1 represents the step wise method mining rare itemsets using hyperlink structure.

Algorithm 1 HLRPM()

```
Input: Complete original transaction_database(DB), rareSup, freqSup
Output: All rare_itemsets in the DB
$\forall$ items in the database, generate the number of times it appears in the DB
for each item, $IT \in DB$ do
  if the support count (IT) $>$ rareSup and support count (IT) $<$ freqSup then
    RareItem $\leftarrow$ IT
  end if
end for each
for each transaction $TR \in DB$ do
  if $\exists$ r | r is Rare $\wedge$ r $\in$ TR then
    RareItemTransaction $\leftarrow$TR
  end if
end for each
Create header table (HT) containing the following fields: item_id, sup_count and hyperlink to
  link the transaction items of TR.
Construct different queues, $Q_j$ that has following fields: item_id and hyper-link that stores the
  items of RareItemTransaction. Use a hyperlink to link all transactions with same first item.
for each item $\beta$ in HT, create rare itemset projections from $\beta$-projected database.
```

The steps in discovering rare itemsets are as follows:

1. The transaction DB is scanned once to find the total number of occurrences of each item in the DB and it is entered in the support table.
2. Initialise two support threshold values, freqSup and rareSup. The itemsets I, having support count between these thresholds are considered to be rare.
3. If the support count of itemsets are lesser than freqSup but greater than rareSup then such items are considered as rare items and it will be included in RareItem table.
4. The DB is scanned again to identify the transactions that contains at least one item that is rare. Only these transactions are included in subsequent steps. Transactions that is a collection of frequent itemsets are excluded from future processing.
5. A header table is created with 3 fields to store item_ID, sup_Count and hyper_Link to link all transactions beginning with the same first item.

6. For each item β in the header table, rare itemset projections will be generated in the new β-projected database.

3.2 Proposed Algorithm

In the proposed algorithm, the header table in [4] is modified such that it consists of following fields: item_id, support_count, Type and hyperlink. Type is an additional field to indicate whether the item is noise(zero), rare(one) or frequent(two). During the initial scan of the database each item is read and its support count is incremented if the item is present in the table. Otherwise, a new entry is created for item in header table if it is not present. The support count is initialised to 1 and Type is set as 0. During each scan of the database, if freqSup > SupCount(I) > rareSup, then its type will be 1 which is referred as Rare item. If SupCount(I) > freqSup then Type is marked as 2 indicating the item is frequent. The additional scan of the database is reduced using this technique. Then rest of the algorithm follows same approach as mentioned in the HLRPM.

Algorithm 2 Efficient Hyper-Linked_Rare_Pattern_Structure (EHLRPS)

Input: Complete original transaction_database, rareSup, freqSup **Output**: Complete set of rare_itemsets

```
Create header table (HT) containing the following fields: item_id, sup_count, Type to indicate
  rare/ frequent and hyperlink
for each item $IT \in DB$ do
  if item IT is in the HT then
    support count(IT)← support count(IT)+1
    if freqSup >support count (IT) > rareSup then
      mark Type (IT) as Rare
    else
      mark Type (IT) as Frequent
    end if
  else if item IT not in the HT then
    create entry for item IT in HT
    support count(IT)← 1
    Type (IT) ← 0
    hyperlink ← NULL
  end if
end for each
for each transaction $TR \in DB$ do
  if ∃ r | r is Rare ∧ r ∈ TR then
    TR ← RareItemTransaction
  end if
end for each
Create different queues, $Q_j$ with following fields: item_id and hyper-link that stores the items
  of $j^{th}$ RareItemTransaction. Use a hyperlink to link all transactions with same first item.
for each item x in HT, create rare item projections from x-projected database.
```

A schematic representation of EHLRPS algorithm is shown in Fig. 1.

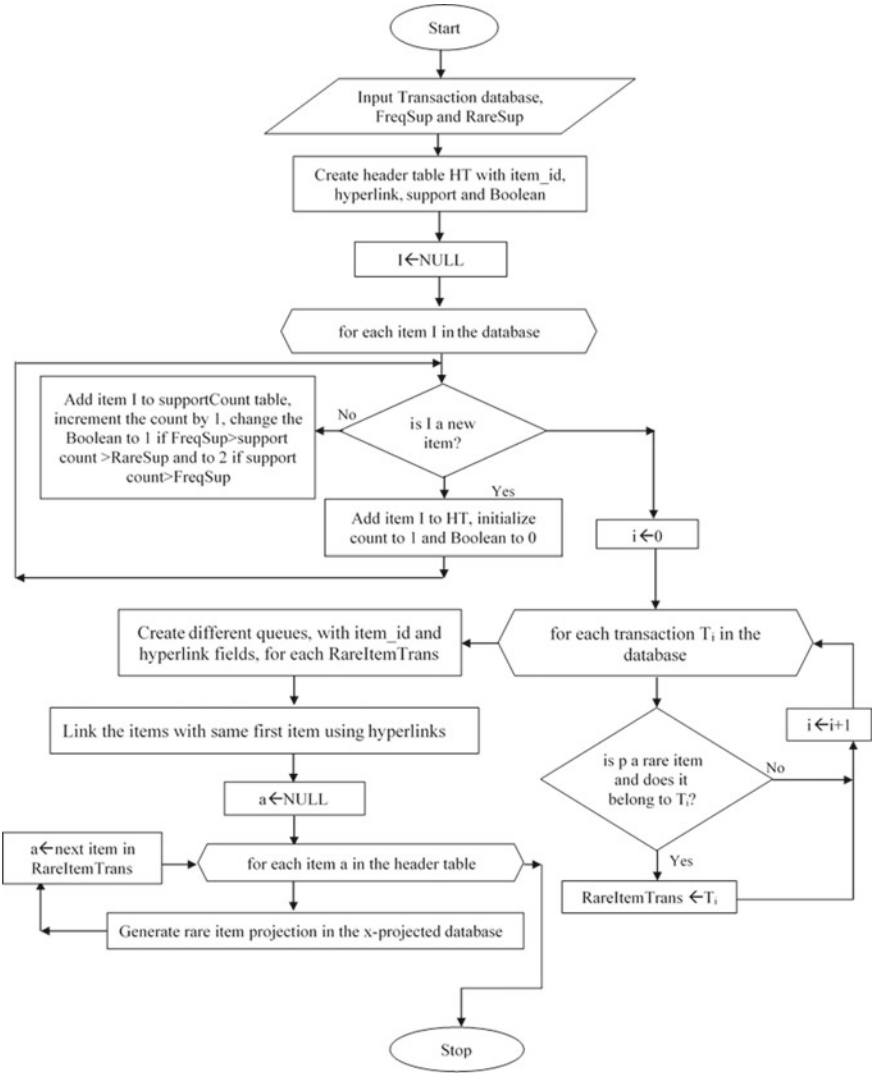

Start
Input Transaction database, FreqSup and RareSup
Create header table HT with item_id, hyperlink, support and Boolean
I←NULL
for each item I in the database
is I a new item?
No
Add item I to supportCount table, increment the count by 1, change the Boolean to 1 if FreqSup>support count >RareSup and to 2 if support count>FreqSup
Yes
Add item I to HT, initialize count to 1 and Boolean to 0
i←0
for each transaction Tᵢ in the database
Create different queues, with item_id and hyperlink fields, for each RareItemTrans
Link the items with same first item using hyperlinks
a←NULL
a←next item in RareItemTrans
for each item a in the header table
Generate rare item projection in the x-projected database
is p a rare item and does it belong to Tᵢ?
No
Yes
RareItemTrans ←Tᵢ
i←i+1
Stop

Fig. 1 Schematic representation of efficient hyper-link RPM algorithm

4 Illustration of the Proposed Algorithm

The implemented method is illustrated by considering a sample database with eight transactions as shown in Table 1.

Table 2 shows the header table HT constructed with four fields, Item_id, Sup_count, Type and Hyperlink, for the transaction database in Table 1.

Table 1 Sample transaction database

T_ID	Itemsets					
Tr1	d	f	c	a		
Tr2	a	d	g			
Tr3	f	e	b	d	a	
Tr4	d	a	h	g		
Tr5	g	a	c	f	e	d
Tr6	d	a				
Tr7	d	a	f			
Tr8	a	e				

Table 2 Header table created based on Table 1

Item_id	Sup_count	Type	Hyperlink
h	1	0	Ø
b	1	0	Ø
c	2	1	Ø
e	3	1	Ø
g	3	1	Ø
f	4	2	Ø
d	7	2	Ø
a	8	2	Ø

Table 3 Header table with rare transactions

Item_id	Sup_Count	Type	Hyperlink						
h	1	0	→	h	g	d	a		
b	1	0	→	b	e	f	d	a	
c	2	1	→	c	f	d	a		
				c	e	g	f	d	a
e	3	1	→	e	a				
g	3	1	→	g	d	a			
f	4	2	Ø						
d	7	2	Ø						
a	8	2	Ø						

Withe respect to the user defined rareSup set to 20% and freqSup set to 40% the rare items identified are **g, e** and **c**. Table 3 shows the output of second database scan that discovers rare transactions.

Table 4 Execution time of hyper-linked RPM and modified algorithm

No. of transactions	Execution time (s)	
	Hyper-linked RPM	Modified algorithm
100	0.545	0.427
200	1.04	0.83
300	1.913	1.339
400	2.159	1.6
500	2.8	2.214
600	3.363	2.446
700	3.683	2.811
800	4.439	3.348
900	4.889	3.811
1000	5.492	4.093

5 Results and Analysis

A comparison between "HyperLinked RPM" algorithm and proposed algorithm is made with data set of 1000 transactions starting with 100 transactions. The experiment was conducted using Visual Studio 2012 on Intel core i5 at 2.30 GHz with 4 GB RAM.

To check the output of sample input transaction database, the algorithm was run with minimum support threshold i.e., rareSup = 20% and maximum support threshold i.e., freqSup = 40%. A total of 10 outputs were analyzed with increment of 100 transactions each time and results were analyzed based on execution time of both the algorithms. Table 4 shows the observed output of both the algorithms. The difference in the time taken for the execution of the algorithms can be noted and using the same values, a line graph representation of the execution time is shown in Fig. 2.

From Fig. 2 it is found that as the number of transactions increases, the modified algorithm is faster and performs 24% better than the "Hyper-Linked Rare Pattern Mining" algorithm.

6 Conclusion and Future Enhancement

The Hyper-Linked Data Structure approach (HLDS) used in discovery of rare itemsets used two tables: Support count table and Header table, which leads to redundancy. This was solved by removing the redundant Support Count table and retaining Header table with an additional field, to indicate the type of the itemset i.e., noise, rare or frequent. The overhead of copying the itemset and its support count to the Header table is also removed which resulted in improvement of the running time.

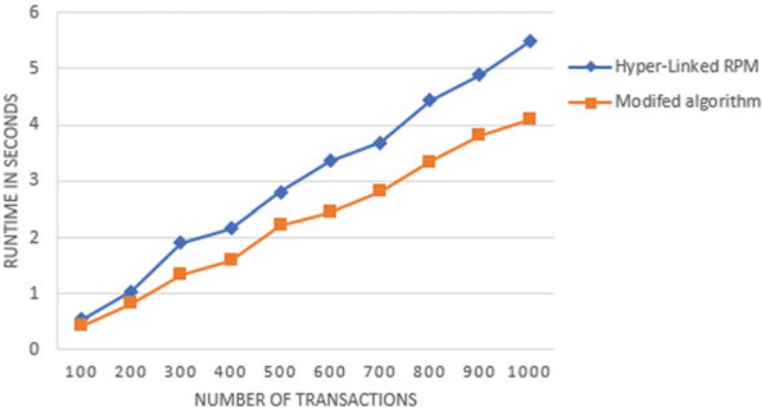


Fig. 2 Execution time of hyper-linked RPM and modified algorithm

The EHLRPS algorithm is a modification of the algorithm given in the paper [4] for mining rare patterns using a queue based HLDS. Performance evaluation, as shown in Fig. 2, explains that the modified method is 24% faster than the algorithm given in the paper [4].

The proposed algorithm works better only when the database is sparse but it is not a good option when the database is dense. It scans the database twice to discover the rare items and then to store the rare transactions in the main memory, a method can be developed to do the same in a single scan as a future enhancement. Also, since the algorithm uses a sequential approach to discover rare itemsets, it is not suitable for big data. Hence a parallel approach to discover the interesting patterns may be implemented.

References

1. Adda M, Wu L, Feng Y (2007) Rare itemset mining. In: Sixth international conference on machine learning and applications (ICMLA 2007). IEEE, pp 73–80
2. Agarwal R, Srikant R, et al (1994) Fast algorithms for mining association rules. In: Proceedings of the 20th VLDB conference, pp. 487–499
3. Bhatt U, Patel P (2015) A novel approach for finding rare items based on multiple minimum support framework. Procedia Comput Sci 57:1088–1095
4. Borah A, Nath B (2017) Mining rare patterns using hyper-linked data structure. In: International conference on pattern recognition and machine intelligence. Springer, pp 467–472
5. Clifton C (2020) Data mining. https://www.britannica.com/technology/data-mining. Last accessed 9 Aug 2020
6. Goyal V, Dawar S, Sureka A (2015) High utility rare itemset mining over transaction databases. In: International workshop on databases in networked information systems. Springer, pp 27–40
7. Han J, Pei J, Yin Y, Mao R (2004) Mining frequent patterns without candidate generation: a frequent-pattern tree approach. Data Mining Knowl Discov 8(1):53–87

8. Kiran RU, Re PK (2009) An improved multiple minimum support based approach to mine rare association rules. In: 2009 IEEE symposium on computational intelligence and data mining. IEEE, pp 340–347
9. Kiran RU, Reddy PK (2009) An improved frequent pattern-growth approach to discover rare association rules. In: KDIR, pp 43–52
10. Kiran RU, Reddy PK (2010) Mining rare association rules in the datasets with widely varying items' frequencies. In: International conference on database Systems for advanced applications. Springer, pp 49–62
11. Kiran RU, Reddy PK (2010) An efficient approach to mine rare association rules using maximum items' support constraints. In: British national conference on databases. Springer, pp 84–95
12. Koh YS, Ravana SD (2016) Unsupervised rare pattern mining: a survey. ACM Trans Knowl Discov Data (TKDD) 10(4):1–29
13. Liu B, Hsu W, Ma Y (1999) Mining association rules with multiple minimum supports. In: Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, pp 337–341
14. Pei J, Han J, Lu H, Nishio S, Tang S, Yang D (2001) H-mine: hyper-structure mining of frequent patterns in large databases. In: Proceedings 2001 IEEE international conference on data mining. IEEE, pp 441–448
15. Pillai J, Vyas O, Muyeba M (2013) Huri–A novel algorithm for mining high utility rare itemsets. In: Advances in computing and information technology. Springer, pp 531–540
16. Szathmary L, Napoli A, Valtchev P (2007) Towards rare itemset mining. In: 19th IEEE international conference on tools with artificial intelligence (ICTAI 2007), vol 1. IEEE, pp 305–312
17. Tsang S, Koh YS, Dobbie G (2011) RP-tree: rare pattern tree mining. In: International conference on data warehousing and knowledge discovery. Springer, pp. 277–288

EDoS-BARRICADE: A Cloud-Centric Approach to Detect, Segregate and Mitigate EDoS Attacks

S. B. Ribin Jones and N. Kumar

Abstract Cloud computing through realizing the height of virtualization offers service models that can meet dynamic demands through performing auto-scaling of resources [1]. This helps the cloud service providers to broaden the grasp across sectors and the computing service market. Though it follows stretchable and elastic service models, it implements a rigid pay-per-use utility pricing model [Ribin Jones and Kumar in J Adv Res Dyn Control Syst 11(9):541–553, 2019 2]. The idea of dynamically scaling across platform makes it more vulnerable to security threats and makes room for easy exploits [Ribin Jones and Kumar in J Adv Res Dyn Control Syst 11(9):541–553, 2019 2]. Among various security threats, the economic denial-of-service (EDoS) attack presents a serious threat, since it exploits auto-scaling feature to impact the utility pricing model [Ribin Jones and Kumar in IEEE Xplore third international conference on trends in electronics and informatics, pp 1003–1008, 2019 3]. In this paper, a real-time cost incurring EDoS attack is performed against a cloud data center hosted Web page with simple Structured Query Language (SQL) manipulation method for experimental research. The experimental observations are applied to define an effective EDoS-BARRICADE that performs detection, segregation and mitigation specific to EDoS attack. The detection algorithm considers metrics that are associated with the auto-scaling feature to detect a suspicious increase in VM activities. The segregation algorithm implements linear SVM to isolate attack VMs optimally and rapidly. The results show that the developed EDoS-BARRICADE algorithms perform detection and segregation with 100% accuracy.

Keywords Cloud computing · Economic denial of service (EDoS) · DDoS · IDPS · SLA

S. B. Ribin Jones (✉) · N. Kumar
Department of Computer Science and Engineering, VISTAS, Chennai, India
e-mail: ribinjones@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_44

1 Introduction

Cloud computing through achieving multi-layer virtualization provides resources for provisioning of on-demand access through Internet access [2]. The resources involve, but not limited to programming platforms, software services and hardware capabilities such as storage, processing and networks [3]. The cloud has the inbuilt brokering capabilities, through which the service users negotiate various characteristics termed as QoS parameters to reach a service-level agreement (SLA) with cloud service providers (CSPs) [3]. Once the agreement is reached, the service offered has advanced and unique features such as on-demand access, pay-per-use, dynamic resource scaling and other characteristics that ensure cloud elasticity [3]. This helps to meet the end-users' varying service demands, with minimal resource wastage. In short, the utility-based pricing model and dynamic resource assignment model of cloud ensures that the user is only paying for the resources used [3]. The auto-scaling feature of cloud helps to realize the on-use resource scaling requirements through auto-allocation or de-allocation of resources; it can scale up or down based on the usage [4, 5]. The auto-scaling at VM level applies to the number of processing units, memory, networking components, etc [6]. Most presumably used auto-scaling metrics from the performance standpoint are threshold and duration [7]. These two metrics are considered necessary for triggering auto-scaling based on the need.

The value-added features of the cloud computing that distinguish it from other distributed and parallel computing model open an unfathomable number of possibilities for exploits and security breaches [8]. Among those, the improvised DDoS attack in the form of EDoS can exploit the cloud pay-per-use model to incur a huge financial loss to the customer or service provider [9]. The evolution of EDoS attack and the limitations in detecting and preventing such attacks has been explained in our previous works [2, 3]. The most threatening of EDoS attack involves exploiting the auto-scaling feature to cause scaling up of VMs in such a way to stealthily evade the security systems and the casual observations. The scaling EDoS attack if persists for a period of time runs VMs with junk data that adds to the offered service and impact heavily on service bill. This can damage the credibility of the service provider and can impact the business heavily and yet goes unnoticed. EDoS attack can adapt botnet model or applies any of the traditional DoS and DDoS attack model on paid services and can impact as EDoS attack.

1.1 Methodology

Cloud computing nowadays is used extensively on Web servers to enhance the request processing potential. This work therefore intents to perform an EDoS attack through iterating SQL-targeted DoS attack. This component does not compromise computers or host any malicious components in server; this helps to avoid security risks and to make sure that the components function without any security hiccups after the completion of the experiment. Moreover, this helps to understand how a simple attack module can constitute a harmful EDoS attack. The obtained results are then used to develop the detection and segregation algorithms that work hand-in-hand to monitor, detect and isolate the compromised VMs from genuine VMs. The Snort rules are then configured to filter and block the isolated VMs. This work improvises the segregation accuracy by incorporating simple and quick linear SVM model. The combined effort is then fitted into a framework to offer a wholesome EDoS handling service.

1.2 Organization

The full article is arranged in the following manner. Section 1 describes the introduction of cloud computing and its security implications from EDoS point of view. Section 2 presents a literature survey of EDoS detection, segregation and mitigation-related works. In Sect. 3, an experimental setting to perform a cloud Web server-targeted EDoS attack is narrated. In Sect. 4, the results of the performed experiment have been analyzed. Section 5 presents the proposed algorithms to detect and segregate attacking VMs or instances using linear SVM. Section 6 discusses the obtained results. Section 7 concludes the article by pointing out the future direction for research.

2 Literature Survey

[1] EDoS-ADS uses the threshold and the duration as auto-scaling parameters, an average CPU utilization threshold with two time triggers scaling-up and down are applied to differentiate VM scaling due to attack from normal. The mode switches to the suspicion mode from normal mode when the cloud CPU utilization exceeds the scaling-up utilization thresholds; then if the attack is found, it switches to attack mode if not it switches to flash crowd mode if the increase continues. In NAT run network, it can block the entire range of attacking nodes from its framework. In [10] EDoS-Shield, the main components of the architecture are virtual firewalls (VF) and verifier cloud nodes (V-nodes). The virtual firewalls are VMs with filtering and routing capabilities that work as filter mechanisms based on the white list to allow authenticated sources and blacklists to block unduly sources. The verifier cloud nodes (V-nodes) are a pool of VM that can verify its legitimacy through Turing tests, and the corresponding white or blacklists are updated based on the results of the verification process. In [4] enhanced EDoS-Shield, as an improvement to EDoS-Shield, it presents two algorithms which enhances both from the perspective of TTL field. Algorithm 1 describes the actions to be taken at VF node upon receiving a packet. A packet can be forwarded to the destination if only its source IP address and TTL value match the white list. Algorithm 2 describes the actions to be taken by a V-Node for not forwarded packets. In [11] EDoS ARMOUR, it is a Multilayered defense Architecture. At the first level, it applies the port hiding mechanism to deter

attackers from knowing the port to perform attacks. At the second level, a learning algorithm is used to monitor user behavior. If inappropriate behavior is spotted, then the corresponding user will get a slower service response. This helps to mitigate application-level DDoS attacks. It also entails a challenge mechanism, admission control, congestion control, user classification and client classification mechanism. Cloud eDDoS mitigation scrubber service in [12] is modeled as an on-demand EDoS-specific security service. Its core functional component implements crypto puzzle to be auto-generated and verified by the users or clients. The users are expected to solve the generated crypto puzzle through brute force to avail the requested service. This helps the CSP to follow the legitimacy of the users. The APART in [8] is a pattern recognition-based EDoS attack mitigation model. The number of packets sent by a user falls in the frequency of 400–800 per second; the APART model gets active and applies pattern recognition to detect EDoS attack. As a deterrence mechanism, it includes time-based and key-sharing post-setup authentication scheme to prevent the replication or replay attacks. It also provides a pre-shared security mechanism to ensure the access of legitimate users on the cloud services. In [9], an enhanced EDoS mitigation mechanism EDOS-EMM has been presented. The design involves three sequential modules, namely (1) data preparation, (2) detection and (3) mitigation modules. Module 1 involves flow monitoring, data collection and processing the flow to segregate it by protocol type and implements the sFlow agent algorithm. Module 2 detects the attack and extracts packet fields such as source, destination IP, port number and no. of packets per second. In this module, Hellinger distance and entropy methods are implemented for anomaly detection to improve accuracy. Module 3 generates alert, initiates rule update process and blocks the attacking IP. In [7] an entropy-based architecture is proposed for the detection of EDoS attacks. The proposed multilayered architecture involves monitoring and aggregation of metrics that affect the cost model, the novelty detection procedures to detect EDoS attack, and the decision-making and action response procedures. It is a predictive mode; it applies entropy variations related with considered metrics such as per-client CPU time when deciding auto-scaling actuations ahead of time for forecasting the EDoS attack possibility.

3 Experimental EDoS Attack

Performing EDoS attack involves components that can corrupt the servers, propagate across the Internet or flood the network or server with traffic with or without the knowledge of the attacker. Hence, performing EDoS attack is not safe in any platform. Therefore, even performing a simple DoS attack over the Internet is prohibited by law. However, without real-time experimentation of EDoS attack, the research cannot be effective. Consequently, an attack which targets the cloud Web server and performs VM scaling, but does not involve compromise, propagation and aftermath exploitation possibility, becomes ideal for our research. However, nowadays, Web servers run in cloud platform to meet any number of request; most of them use pay-per-hit model [13]. Therefore, if the Web servers are forced to process more requests, then it starts more VMs and ends up as a complicated EDoS attack. This requires writing a Web page and hosting it in a cloud Web server, and to use a minimal number of client computers to start consecutive DoS attacks together, they cause an EDoS attack. This way, single client computer can start any number of VMs in the server that stays and performs EDoS attack on server without compromising the facility.

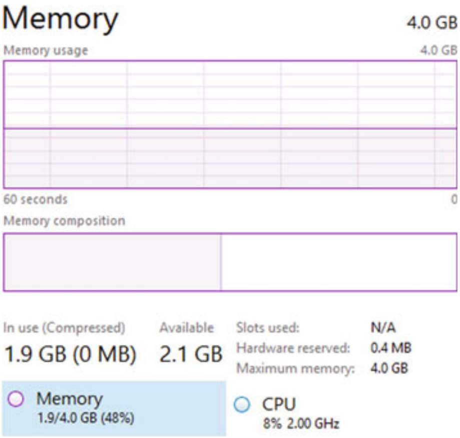
The objective of this section is to perform a real-time EDoS attack on a Web hosting data center. The experimental setup involves a Windows Server 2016 and 25 machines connected through a LAN. The server runs in a cloud platform. The experimental scenario involves setting up a client Web page and creating db tables in server for Web access. Initially, the database for Web application is loaded in the server with various user names. This helps to ensure that none of the server which runs Web pages is targeted by the attack. The client sends the request; the server uses VMs to create a separate instance to process every request. Usually, once the request is complete processing, the VM is released. However, to perform database-targeted EDoS attack, it involves three steps. Step 1: The blah application is used to insert a user name and password from the client browser. Blah is a database manipulation command which helps to perform SQL injections and DoS attacks [14]. Step 2: To start a DoS attack, a 'blah' shell command is delivered to start a separate VM for every request as well as to performs a half-open TCP SYN-based DoS attack. The command **blah'; exec master.xp_cmdshell 'ping** target Web page **-165000 –t'**; – create a VM in server-side and that VM acts on behalf to reply to the client. It then performs a half-open TCP ping-based DoS attack. Step 3: The attack performs VM scaling by starting five VMs every 5-min interval from five computers through repeating step 2 for 5 times. This avails five VMs every interval. The cost of availing every VM for one minute is set to 0.0011 [8]. The hypervisor access is prohibited to avoid single-point failure because if it fails, the entire cloud facility becomes inaccessible. The VM or workloads of the experiment is monitored by task manager and Wireshark applications.

4 Experimental Results of EDoS Attack and Observation

The Windows Server on which the EDoS attack is performed is hosting the cloud-based file sharing application termed as 'ownCloud.' ownCloud becomes inaccessible due to the attack from the local client. Since every VM responds individually on behalf of the server, it becomes difficult for the firewalls to detect. The experiment has been conducted in a confined environment only after making sure that the Windows Defender Firewall and Web Shield are turned ON. However, they failed to detect the attack due to two reasons; one is the projected traffic which happens with in the server, that is, VM or frontal instances and Web server runs within the server due to the cloud incorporation. The firewall that looks for anomaly from the boundary of the server, therefore, fails to detect. Another reason is that the VMs send traffic in an individualized manner, so collectively, it will not shoot up at firewall observation point. Therefore, the firewalls fail to register the flood.

The results are observed from the perspective of traffic, processing and memory. According to Figs. 1 and 2, it is evident that the temporary memory usage is increasing gradually, i.e., from 1.9 (48%) to 2.1 (53%). Initiation of application involves starting and allotting space for VMs; therefore, more memory usage is evident in Fig. 1. However, in Fig. 2, due to the similarity of data processed and traffic generated, it did not reflect on memory or processing even after starting additional 20 frontal instances. However, for genuine application, each VM may process unique tasks, so it consumes more processing capability and reflects upon processing and memory.

The Web server due to the cloud implementation allows any number of frontal instances or VMs to be deployed by a single click from the client computer. It also allows every VM to act as the server to the client. At the same time, it manages to receive tremendous TCP hits that are generated from the VMs. In spite of all this, the server functions smoothly except for the subordinate application such as ownCloud.



Memory 4.0 GB

Memory usage 4.0 GB

60 seconds 0

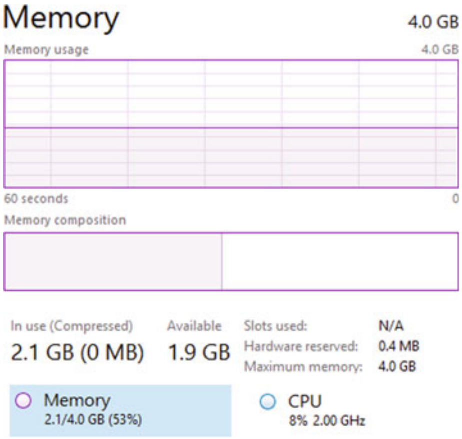
Memory composition

In use (Compressed): 1.9 GB (0 MB)
Available: 2.1 GB
Slots used: N/A
Hardware reserved: 0.4 MB
Maximum memory: 4.0 GB

Memory 1.9/4.0 GB (48%)

CPU 8% 2.00 GHz

Fig. 1 Memory observation after starting of five VMs



Memory 4.0 GB

Memory usage 4.0 GB

60 seconds 0

Memory composition

In use (Compressed): 2.1 GB (0 MB)
Available: 1.9 GB
Slots used: N/A
Hardware reserved: 0.4 MB
Maximum memory: 4.0 GB

Memory 2.1/4.0 GB (53%)

CPU 8% 2.00 GHz

Fig. 2 Memory observation after starting of 25 VMs

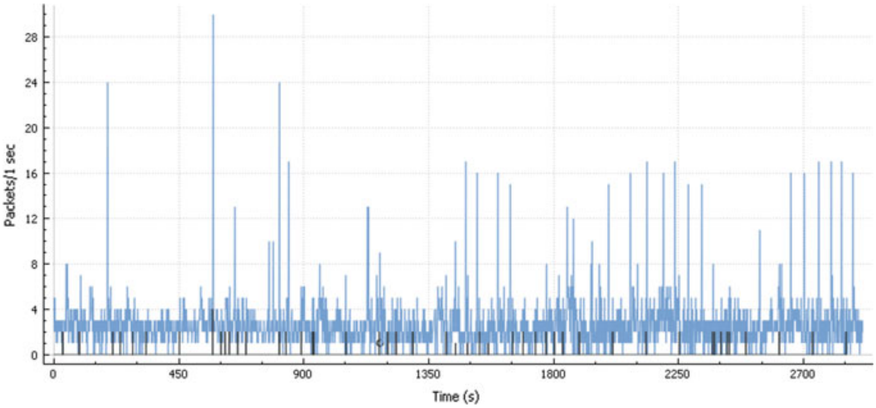


Fig. 3 Horizontal increase of incoming TCP packets from 25 VMs toward the server

This is because the Web service is prioritized over the ownCloud sharing application. The traffic anomaly is however observed with Wireshark, and the observation is presented as the following graph.

According to Fig. 3, the TCP traffic increases in horizontal manners in specific to the VMs, which is uncustomary for a flood attack. Moreover, occasional hits from the database (DB) or Web server were also visible from the graph, denoted as black bars. Consequently, the results show the requirement for individualized monitoring of VMs both from the perspective of the following auto-scaling metrics and behavior of VMs. Based on the experimental results, the following detection and segregation algorithms were designed.

5 Proposed EDoS-Barricade

The complications of cloud security arise from its unique features. The proposed EDoS-Barricade is therefore designed to look into the cloud features closely to eliminate the most threatening form of cloud attack, i.e., EDoS. Evidently after close examination of the experimental results, for initial EDoS detection, following the number of VMs with respect to a predetermined interval becomes necessary to deter the auto-scaling-based attack from exploiting the cloud pay-per-use model. The detection algorithm, therefore, requires to follow the VM to server communication and capturing to-and-fro packets. However, to achieve that, the Snort firewall has been chosen. The Snort is a widely used firewall for Linux, and it suits both research and commercial purposes and helps to build OS kernel-level solution for security implementation. However, Snort is tedious to work with, after careful tuning of Snort SO rules, pre-processor rules, etc.; it still lacks the proper command to detect VM to server traffic. After careful inspection, the following command is tailored to detect VM to server communication; so far, there is no reference to do that .

Table 1 Snort VM to server detection procedure

Innovative Snort command to detect in-server VM to server hits
alert tcp $HOME_NET any any -> $HOME_NET any flow:to_server; Seq: N; ack: A;
win: W;length: l; Options:O; classtype:attempted-user;

The above command helps to detect VMs with its IDs, as well as to follow all sorts of in-house traffic between VM and server communication; it can also follow server to VM communications.

5.1 Auto-Scaling-Based EDoS Detection

The 15 number of VMs is identified as the reasonable number for presetting detection metric from our experimentation standpoint. The chosen interval is 5 min. This number can be deduced by looking into cloud history. For instance, the average number of VMs running with respect to the time interval for a period of history, such as for a month and the year, can be computed to set this parameter. If the number of VMs increases for three consecutive intervals, then this algorithm raises the detection alert as well as switches mode to suspicious and calls the EDoS-BARRICADE Segregation () algorithm.

5.2 Support Vector Machine (SVM)

The EDoS segregation requires classifying data into two binary sets {EDoS, normal}. However, to achieve that, the SVM classifier is an ideal choice. SVM offers both linear and nonlinear mapping techniques to project and classify the dataset or input into 2-D or 3-D plane. This projection helps to deduce a hyperplane that can segregate data into its corresponding classes. The linear SVM can offer quick and accurate classification compared to other classifiers. Moreover, SVM can stretch to big data processing since it can handle outlier and overfitting effectively. The dataset during the preprocessing is usually depicted as follows;

$$(x_1,\ y_1),\ (x_2,\ y_2),\ldots(x_i,\ y_i),\ldots(x_n,\ y_n) \tag{1}$$

where x_i denotes the metrics that can help to classify the given data; from our observation, they are incoming and outgoing packets in specific to VMs, n denotes the maximum number of values, and i denotes the iteration of values from 1 to n.

$$\text{Class}x_i \in \{\text{Incoming\_pkts},\ \text{Outgoing\_pkts}\} \tag{2}$$

The class y_i serves as a label and helps to separate the VM data into its corresponding class.

$$\text{class}\, y_i \in \{\text{EDoS,normal}\} \tag{3}$$

Given the datasets x_i and $y_{i,}$ the SVM classification engine naturally creates a hyperplane as follows

$$w^{\mathrm{T}} * x + b = 0 \tag{4}$$

The hyperplane is a separator that determines the maximum margin between two classes, where w is a weight vector and b is a bias. The SVM automatically creates weight vector and bias to group the data using a hyperplane. The grouping is done as follows.

$$\text{if} \quad w^{\mathrm{T}} x + b \geq +1 \text{ then normal} \tag{5}$$

$$\text{if} \quad w^{\mathrm{T}} x + b \leq -1 \text{ then attack} \tag{6}$$

However, the closest point called a support vector from the respective classes can only establish the optimally separating hyperplane. Support vector from the given dataset determines the hyperplane by using:

$$\min(\tau) = \left\{{}^{1}\!/_{2}||w||^{2}\right\} \tag{7}$$

It has been shown that the optimal separating hyperplane can be found by minimizing (7), with respective to

$$y_i(w,\ x_i + b) \quad \text{for all} \quad i = 1, \ldots, n \tag{8}$$

This is a convex optimization problem; the SVM classifier solves it with Lagrange multipliers

Algorithm:1 EDoS-BARRICADE Detection ()

```
Step 1: Monitor increase or decrease in workloads or VM scaling
Step 2: If no. of VM increases above the default limit then initiate EDoS detection
Step 3: If it continues to increase for consecutive intervals, raise detection alert and spontaneously call the segregation procedure
```

(continued)

(continued)

Algorithm:1 EDoS-BARRICADE Detection ()

```
Input: {List of VMs with unique identifier, Interval}
Output: {List of VMs, No. of, Interval, Total No. of incoming & Outgoing pkts}
Define VM_Number; \\ Number of workloads
Define interval = 5 min; \\ Time Interval
Define i = 0; \\Increment Variable
For each application;
if VM_number > 15 then Monitor;
For each interval;
if current_VM_No > previous_VM_No;
{ i +1
if i > 3 then raise detection alert
Switch Mode to Suspicious
Call EDoS-BARRICADE Segregation () }
```

5.2.1 Linear SVM

The dataset contains the non-separable case of data; however, by introducing a cost for violating constraint (8) such as a positive slack variable ξ, a hyperplane can be realized for optimal separation.

$$y_i.(x_i.w + b) \geq 1 - \xi_i. \quad \text{For all } i \tag{9}$$

It can be furthered for erroneous cases with cost parameter C, and such a scenario is not required for this research.

5.2.2 Nonlinear SVM

If the data is not mapable using linear decision function, the data will be mapped into higher dimensional feature space through a nonlinear transformation which incorporates kernel functions. Vastly used kernel functions for binary classification problems are as follows;

$K(x_i, x_j) = x_i x_j$: *linear SVM*
$K(x_i, x_j) = (x_i x_j + 1)^p$: *polynomial of degree p*
$K(x_i, x_j) = \tanh(a.x_i x_j + b)$ *and* $a > b$: *multi-layer perceptron (MLP) classifier*
$K(x_i, x_j) = \exp\{||x_i - x_j|^2/^2\}$: *radial basis function (RBF) classifier.*

5.3 Behavior Heuristics-Based EDoS Segregation

The training dataset has been generated during the monitoring with respect to VMs. The data is provided to an SVM-based EDoS-BARRICADE segregation algorithm to differentiate malicious VMs from normal VMs.

Algorithm:2 EDoS-BARRICADE Segregation ()

```
Step 1: Upon invocation from the detection algorithm; Gather the data.
Step 3: Perform anomaly-based segregation by applying relevant SVM.

Input : {List of VMs, No. of, Interval, Total No. of incoming & Outgoing pkts}
Output: {Scatter plot, Attacking VM list for snort to filter and block }
Call SVM Classifier ()
For each VMs;
{ if $wT x + b \geq + 1$
then project as normal
if $wT x + b \leq -1$
then project as EDoS }
For all VMs;
{if linearly separable
Deduce hyperplane $yi.(xi.w + b) \geq 1 - \xi i$
else apply kernel}
project result in scatter plot
Generate Attack VMs list
Call Snort()
Perform Filtering and Blocking
```

The above algorithm presents the result for input dataset instantaneously; it can be plotted into two classes separated with a hyperplane. Moreover, the output presents a list of attacking VMs which can be blocked and filtered out. The following sections examine the result of EDoS-BARRICADE through varying the input.

6 EDoS-Barricade Result Observation and Verification

The experimental data along with normal VM dataset obtained from GitHub has been fed into RapidMiner. The linear SVM has been simple enough to classify the normal flow from the attack flow as follows.

According to Fig. 4, attacking VMs distinguishes itself with the following heuristics, i.e., for attacks, ‘Accumulated incoming packet > Accumulated normal Incoming packets’ and ‘Accumulated Outgoing packet < Accumulated normal Incoming packets.’ Moreover, to intensify the testing, data migration dataset obtained from GitHub has been incorporated into the dataset, and then, testing is performed. For that, the SVM is required to transform data into a higher dimensional plane using a kernel as shown in Fig. 5.

According to Fig. 5, the attack instances become as blue dots becomes classifiable using SVM kernel when the data migration dataset is involved. However, this is an

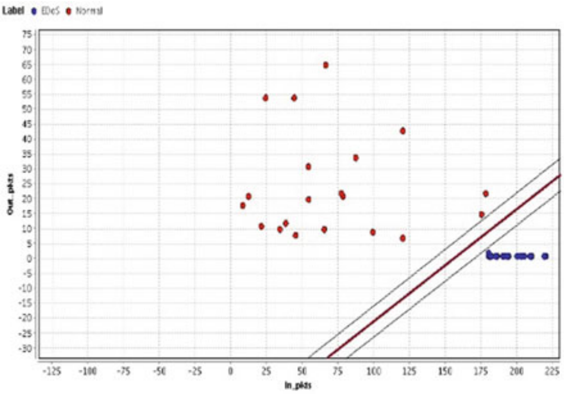


Fig. 4 Linear SVM segregation result after the third interval

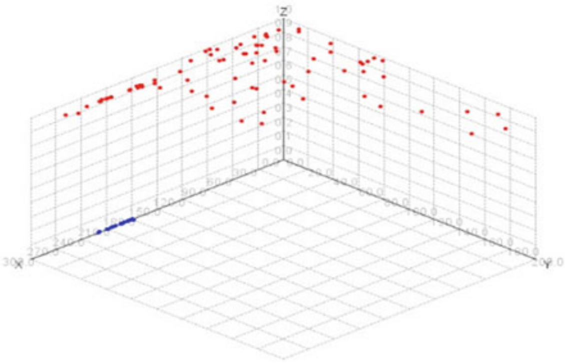


Fig. 5 Nonlinear SVM segregation result for a dataset with data migration instances

unusual and rare happening; for other cases, simply linear SVM proves sufficient to detect EDoS attack. The accuracy achieved for EDoS detection after the third interval is 100%. The obtained SVM-based EDoS-BARRICADE result is then interpreted as the list of attacking VMs and forwarded to the Snort firewall for filtering and blocking.

7 Conclusion and Future Scope

The EDoS attack is performed on any cloud Platform through exploiting the auto-scaling feature to increase the number of VMs, which in reality processes junk data and floods the server with junk requests. This phenomenon manages to exploit the cloud usage-based pricing model and impact heavily on customer bills. However, to have a real-time perspective, a Web server-based EDoS attack is performed, and the impact of the attack is observed. The result shows that the EDoS attack is detectable by following the auto-scaling and segregate-able by following incoming and outgoing packets. In accordance with the results and through making necessary improvement to the Snort, a quick auto-scale-based EDoS-BARRICADE detection algorithm has been proposed. The tedious part involves isolating attack VMs from the normal ones without false detections. An EDoS-BARRICADE segregation algorithm which involves linear SVM as its base classifier is proposed and tested with datasets involving normal and attack data. The linear SVM performs classification with 100% accuracy since the normal data has more number of incoming packets and less number of outgoing packets than the attack data. However, more rigorous testing has been performed by including data migration samples to the dataset. In such cases, the algorithm switches to nonlinear kernel-based SVM classifier to isolate the attack class in higher dimensional plane. In the future, the presented work will be extended into a framework with more added characteristics.

References

1. Shawahna A, Abu-Amara M, Mahmoud ASH, Osais Y (2018) EDoS-ADS: an enhanced mitigation technique against economic denial of sustainability EDoS attacks. IEEE Trans Cloud Comput (Feb)
2. Ribin Jones SB, Kumar N (2019) Unraveling the security pitfalls that stem from core cloud benefits through analyzing various DoS attacks, detection and prevention. J Adv Res Dyn Control Syst 11(09):541–553 (Aug)
3. Ribin Jones SB, Kumar N (2019) Precursory study on varieties of DDoS attacks and its implications in Cloud Systems. In: IEEE Xplore third international conference on trends in electronics and informatics, pp 1003–1008, Apr 2019
4. Sqalli MH, Al-Haidari F, Salah K (2012) Enhanced EDOS shield for mitigating EDoS attacks originating from spoofed IP addresses. In: IEEE eleventh international conference on trust, security and privacy in computing and communications
5. Kumar D (2019) Review on task scheduling in ubiquitous clouds. J ISMAC 1(01):72–80
6. Bulla S, Basaveswara Rao B, Gangadhara Rao K, Chandan K (2018) An experimental evaluation of the impact of the EDoS attacks against cloud computing services using AWS. Int J Eng Technol 7(1.5):202–208
7. Monge MAS, Vidal JM, Villalba LJG (2017) Entropy-based economic denial of sustainability detection. MDPI Entropy 19(12) (Nov)
8. Thaper R, Verma A (2015) Adaptive pattern attack recognition technique (APART) against EDoS attacks in cloud computing. In: IEEE third international conference on image information processing
9. Singh P, Rehman SU, Manickam S (2017) Enhanced mechanism to detect and mitigate economic denial of sustainability (EDoS) attack in cloud computing environments. Int J Adv Comput Sci Appl 8(9)
10. Sqalli MH, Al-Haidari F, Salah K (2011) EDOS shield—a two-step mitigation technique against EDoS attacks in cloud computing. In: IEEE fourth international conference on utility and cloud computing
11. Masood M, Anwar Z, Raza SA, Hur MA (2013) EDoS Armor: a cost effective economic denial of sustainability attack mitigation framework for e-commerce applications in cloud environments. In: IEEE sixteenth international multitopic conference
12. Kumar MN, Sujatha P, Kalva V, Nagori R, Katukojwala AK, Kumar M (2012) Mitigating economic denial of sustainability (EDoS) in cloud computing using in-cloud scrubber service. In: Fourth international conference on computational intelligence and communications networks

13. Abusitta A, Bellaiche M, Dagenais M (2018) An SVM-based framework for detecting DoS attacks in virtualized clouds under changing environment. Springer J Cloud Comput: Adv Syst Appl 7(9):1–18 (April)
14. Mazrekaj A, Shabani I, Sejdiu B (2016) Pricing schemes in cloud computing: an overview. Int J Adv Comput Sci Appl 7(2):80–86
15. Duraipandian M, Vinothkanna R (2019) Cloud based internet of things for smart connected objects. J ISMAC 1(02):111–119
16. Udhayan J, Hamsapriya T, Vasanthi NA (2012) DDoS attack detection through flow analysis and traffic modeling. In: SPIT 2011, LNICST 62, pp 89–94
17. Baig ZA, Sait SM, Binbeshr F (2016) Controlled access to cloud resources for mitigating economic denial of sustainability (EDoS) attacks. Elsevier Comput Netw 97:31–47 (Mar)

User Query-Based Automatic Text Summarization of Web Documents Using Ontology

K. Selvakumar and L. Sairamesh

Abstract Web document summarization is a very important function for knowledge management when documents are very huge and dynamic. This helps the document readers to easily read and understand. In general, summarization is done based on the extraction of sentences from the retrieved documents and assembled with the needed proposition. Text content plays a major role in the Web, and extraction of text content is a challenging task. Recently, one of the problems that arise with the rapid growth of the Web and general information availability (sometimes referred to as an information overloading) is the increased need for effective and powerful text summarization. In this paper, a system is built for automatic text summarization from the source documents which is retrieved from the Web. The proposed method considers the ontology approach to extract the text summary according to query terms, which are existing in the ontology graph based on their depth. Also, attributes are used to improve the semantic representation of a sentence's information from contents.

Keywords Information retrieval (IR) · Text summarization · Ontology · Automatic text summarization (ATS)

1 Introduction

Automatic text summarization (ATS) is a precise version of a text provided by a computer program [1]. Normally, summarization is the one that summarizes the content of the document without changing the actual meaning of the document. In this task, summarizing the content with the same synonyms is a challenging task. To solve this challenge, text summarization tries to summarize the text content which

---

K. Selvakumar (✉)
Department of Computer Applications, National Institute of Technology, Trichy, India
e-mail: kselvakumar@nitt.edu

L. Sairamesh
Department of IST, CEG Campus, Anna University, Chennai, India
e-mail: sairamesh.ist@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_45

has to provide brief information on the content. Some of the tools are available for summarization which searches the headings and subtopics relevant to the key points given in the document and gives the content [2, 3].

Mostly, text summarization can be classified into two different ways. One is the extractive method which consists of important sentences from the given paragraph. It extracts the important sentence based on the frequency of words and linguistic features of the word. The next is abstractive summarization, which is mainly based on language concepts or linguistic features of the context. It uses the natural language processing approach to examine the concepts and express them in the shortest context than the available version. In this paper, the context is summarized based on extractive summarization.

In the operation of extractive summarization, the key segments are formulated based on a statistical analysis of the text. The features of sentences are extracted based on word frequency and location of words. The frequent words are taken for the context by treated as the most important word in the context. This type of extraction makes the user to easily understand the concept without having deep knowledge of the topic. The process of extractive summarization is classified as sentence boundary identification, stop word elimination, stemming, tagging, frequent word identification, and scoring of sentence [4, 5].

In the processing, features of the sentences are computed and selected based on the influence of the words in the sentences, and weights are assigned based on the context features. The ranking is done by using the feature—weighting equation and sentences chosen for the final summary.

Problems with the existing techniques for generating extractive summary are:

(a) The sentence which is ranked high may also contain some unwanted information.
(b) Relevant information is available across the sentences, and hence, the extractive summaries will not be able to capture this sometimes due to the length of the sentence.
(c) Sometimes, conflicting information may not be presented precisely.

Interpretation of anaphors among similar sentences may lead to include some unwanted information in summary. Similar challenges are present with time-based expressions. These problems most commonly arise in multi-document summarization, since information is collected from different sources. In such a scenario, post-processing can be used to extract anaphors by replacing the pronouns with their antecedents by replacing the relative temporal expressions represented using dates.

2 Current Practice and Research

Many works are proposed by most of the linguistic and information retrieval researchers for multi-document summarization. The work proposed in [4] used A* search technique with discriminative training approach for summarizing the multiple documents. In this approach, A* is used to search and provide the relevant words for the given keywords and make the summary relevant to the query. In [6, 5], multi-document summarization is done by using the ontology which easily extracts the meaning of the keyword and finds the appropriate words for the keyword and summarizes the content. Some other works in [7–9] explain various summarizations for the single and multi-document. In [10], the author describes how to evolve the user groups which easily evaluate a summary for the group of people. In [11], summarizations are carried out based on the relationship between the contexts using natural language processing.

3 Ontology Approach in Text Summarization

The usage of ontology is increased due to the increase in word usage in different ways. Gruber's define "ontology as an explicit specification of the topics." In general, ontology is a formal representative of the vocabulary based on the logical definition of the words and statements where it is used. It provides a formal way to easily get the relevant information for the given keywords. Assume that large articles contain many subtopics within the main topic. So, the keywords available in the subtopic have to be identified, and it makes the summarization easier and meaningful.

Figure 1 shows the architecture of the text summarization system proposed in this paper. In this proposed method, user query will be preprocessed (such as stop word removal and stemming), and key terms are extracted. Each term is compared with the existing ontology graph, and weight will be assigned according to the depth where they exist in the ontology graph [12]. Here, the ontology graph can be constructed using the protégé tool. The key terms which are present in the ontology graph will be estimated and ranked; otherwise, the value of those terms will be assigned as -1. Based on the rank, the key terms are chosen, and their corresponding statements will be extracted from various multiple Web documents. Before this step, the Web documents are preprocessed. After extracting the sentences from the documents, their similarity can be measured by using similarity measures (such as cosine similarity). The above process can be applied to specific sports domain areas.

Let A be the selected features. For each candidate document x_i, its dynamic features used for scoring (Eq. 1) include the similarity measure between the sentence and user query.

$$\text{Score}(x_i) = \frac{(\Psi \text{Sim}_1(x_i, Q) + (1 - \Psi)\text{Sim}_2(x_i, A))n(x_i)}{N} \tag{1}$$

where

N	Total number of documents.
$n\,(x_i)$	Number of sentences having similarity measure above the predefined threshold level.

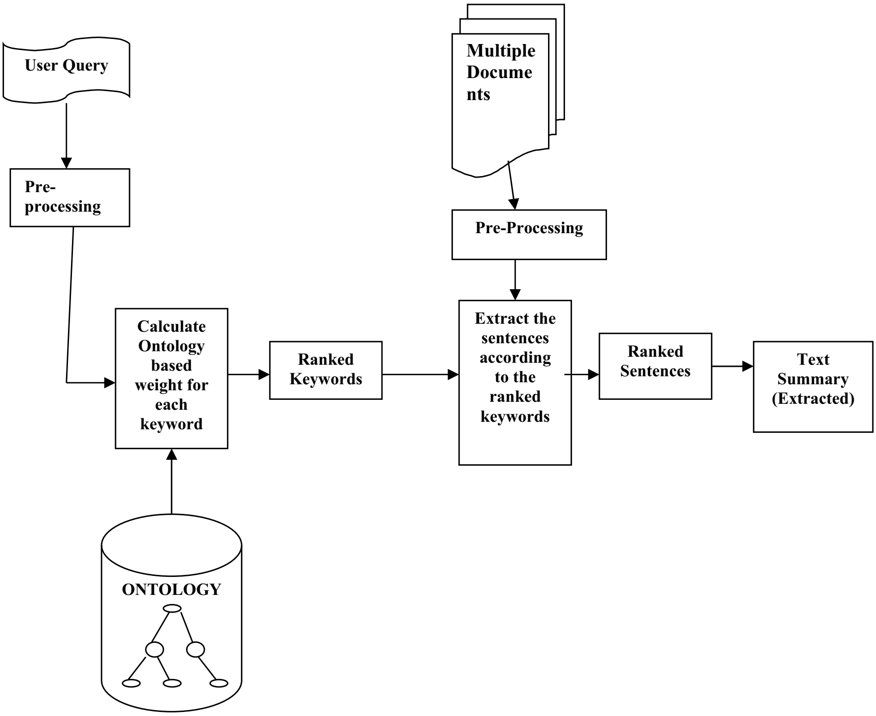


Fig. 1 System architecture for text summarization

Ψ Parameter with values [0, 1] that control relative importance given to relevance versus redundancy.

Sim_1 and Sim_2 are similarity measures (Eq. 2) with documents (x, y) selected feature (research area) and user query.

$$\mathrm{Sim}_1(x, y) = \mathrm{Sim}_2(x, y) = \frac{\langle x, y\rangle}{\|x\| \cdot \|y\|} \tag{2}$$

To represent the structural information of the words and their taxonomy, it propagates feature constraints from the leaf nodes to the parent nodes by recursively aggregating the children feature weights and assigning them to the parent:

$$\omega_c'(t) = \omega_c(t) + \sum_{i \in \mathrm{children}(c)} \omega_i(t) \tag{3}$$

where

$\omega_i\ (t)$ is the TFIDF weight of term t in child category i.

4 Results and Discussions

The results for text summarization are carried out with the different set of documents which are varied in size (MB). Table 1 shows the summary of the data set which is collected from newspaper sites under different topics such as sociology, sports, education, entertainment, and medicine.

The content of the document is taken as the average sentence, and experiments are executed, and the results are compared with the existing systems. Table 2 shows the summary accuracy of the proposed and existing systems. The proposed system shows better accuracy than the available system for summarization. Moreover, the result shows that the proposed system performs better both semantically and also in content summarization than [2] which follows latent semantic analysis and [13] uses hierarchical summarization for multiple documents. Table 3 shows the comparison of the summarized content of the existing and proposed systems. The proposed system summarizes the content of different documents as a single document and provides a semantically acceptable summarization.

Both the accuracy and summarization are better proposed than the existing system. And also, the time taken for summarization is calculated for all experimental systems as shown in Table 4, where most of the researchers not concentrated on the timing constraints in summarization. In this manner, our proposed system is performing better than the existing approaches.

Table 1 Data set for summarization

Data set (DS)	Number of documents	Average number of sentences per document
DS1	35	1230
DS2	58	1080
DS3	75	950
DS4	63	750
DS5	70	930

Table 2 Summary accuracy

Data set	Existing system [2] (%)	Existing system [13] (%)	Proposed system (%)
DS1	84	92.4	96.5
DS2	90.8	93.5	97.4
DS3	92.3	94.7	96
DS4	87.5	91	97.8
DS5	91.5	95	98.4

Table 3 Comparison of summarized content

Data set	Avg. no. of sentences per document	No. of sentences in summarized content		
		[2]	[13]	Proposed system
DS1	1230	1750	1680	1500
DS2	1080	980	1100	930
DS3	950	2360	2180	1800
DS4	750	1950	1800	1730
DS5	930	1680	1550	1420

Table 4 Time is taken for summarization

Data set	Existing system (ms) [2]	Existing system (ms) [13]	Proposed system (ms)
DS1	21.4	18	9.5
DS2	15	16.5	8.4
DS3	57.5	56.3	14.8
DS4	54	48	13
DS5	46	35	12.6

5 Conclusion and Future Work

This user query-based automatic text summarization is an old challenge, but the current research direction leans toward emerging trends in biomedicine, product review, education domains, emails, and blogs. This is because there is information overload in these areas, especially on the World Wide Web. This paper examines recent advances and challenges of automatic text summarization in general and explores some emerging trends on automatic text summarization using the ontology approach. Pertinent issues persist in automatic text summarization, especially that of achieving summarizations that are close to those produced by human linguists. However, even expert summarizations have slight differences. Future work includes extractive summarization using a fuzzy neural network approach to enrich its efficiency.

References

1. Mitkov R (2005) The Oxford handbook of computational linguistics. Oxford University Press
2. Ozsoy MG, Cicekli I, Alpaslan FN (2010) Text summarization of Turkish texts using latent semantic analysis. In: Proceedings of the 23rd international conference on computational linguistics, 2010. Association for Computational Linguistics, pp 869–876
3. Ramesh LS, Ganapathy S, Bhuvaneshwari R, Kulothungan K, Pandiyaraju V, Kannan A (2015) Prediction of user interests for providing relevant information using relevance feedback and re-ranking. Int J Intell Inf Technol (IJIIT) 11(4):55–71

4. Aker A, Cohn T, Gaizauskas R (2010) Multi-document summarization using A* search and discriminative training. In: Proceedings of the 2010 conference on empirical methods in natural language processing. Association for Computational Linguistics, pp 482–491
5. Li L, Wang D, Shen C, Li T (2010) Ontology-enriched multi-document summarization in disaster management. In: Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval. ACM, pp 819–820
6. Guran A, Bekar E, Akyokus S (2010) A comparison of feature and semantic-based summarization algorithms for turkish international symposium on innovations in intelligent systems and applications. Kayseri Cappadocia, Turkey, pp 21–24
7. Litvak M, Last M, Kisilevich S, Keim D, Lipman H, Gur AB (2010) Towards multi-lingual summarization: a comparative analysis of sentence extraction methods on English and Hebrew corpora. In: Proceedings of the 4th workshop on cross-lingual information access, pp 61–69
8. Naderi N, Witte R (2010) Ontology-based extraction and summarization of protein mutation impact information. In: Proceedings of the 2010 workshop on biomedical natural language processing. Association for Computational Linguistics, pp 128–129
9. Foong OM, Oxley A, Sulaiman S (2010) Challenges, and trends of automatic text summarization. Int J Inf Telecommun Technol 1(1)
10. Selvakumar K, Ramesh LS, Kannan A (2015) Enhanced K-means clustering algorithm for evolving user groups. Indian J Sci Technol 8(24):1
11. Chakrabarti P, Basu JK (2010) Text summarization, and discovery of frames and relationship from natural language text-A R&D methodology. Int J Comput Sci Eng 2(3):487–492
12. Wan X, Yang J (2008) Multi-document summarization using cluster-based link analysis. In: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval. ACM, pp 299–306
13. Christensen J, Soderland S, Bansal G (2014) Hierarchical summarization: scaling up multi-document summarization. In: Proceedings of the 52nd annual meeting of the association for computational linguistics, pp 902–912

Quantum Chaos-Based Encryption Technique for Transmission of Medical Images

R. Anitha and B. Vijayalakshmi

Abstract With the fast advancement of networked devices and growth in usage of digital data transmission, hawk-eye attention is needed to provide the secured transmission for various vital data such as medical images and digital signatures. Even though many encryption algorithms were proposed, achieving high-security transmission remains in the darker side of the research because of its low sensitivity and less complexity. Recently, usage of quantum chaos has gained more insight in the encryption world due to its complex nature. A new encryption scheme quantum chaos-based network-centric encryption for data transmission (Q-CNEST) is proposed to integrate with the network characteristics as its initial conditions for generation of high complex encrypted data. As a first step, cipher keys are generated by the quantum chaotic logistic maps using the random network parameters. Following that, diffusion and permutation of each pixel with the high randomness cipher keys have been performed. Finally, sensitivity analysis is done. Simulation has been implemented in the Qiskit packages, and its results have proven that the proposed algorithm has good randomness which will be more suitable for the prevention of conventional attacks over the medical information. Moreover, the proposed encryption algorithms can provide more limelight for the secure transmission of medical image transmissions.

Keywords Cipher keys · Qiskit packages · Digital imaging and communication in medicine (DICOM) · 2D spatiotemporal chaotic-based encryption · Internet of things (IoT) · Long-term evaluation (LTE) · Received signal strength (RSSI)

1 Introduction

In the current scenario, most of the industrial applications are highly associated with the revolution of technology. The advancements in technology have revolutionized medical field and telemedicine. The security of patient health record is imperative

---

R. Anitha (✉) · B. Vijayalakshmi
B.S.Abdur Rahman Crescent Institute of Science & Technology, Chennai, India
e-mail: anitharajesh29@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_46

from security threats. Digital imaging and communication in medicine (DICOM) is a standard that focuses on security issues of telemedicine. It insists on security services like privacy, reliability and authenticity. The secrecy of transmitted images is achieved by symmetric encryption algorithms [1], whereas reliability and authenticity are achieved by hashing and digital signatures. But all these classical encryption techniques encounter a complex computation. Hence, the development of an efficient algorithm has turned out to be an interesting research field. Chaos theory has many attractive features, like sensitivity to initial values, pseudorandomness and ergodicity and low cost in the computer operating system. Hence, a chaotic system suite best for encryption process. In recent years, chaos-based image encryption algorithms have become an interesting research domain.

However, the current chaotic system still needs improvisation in terms of high randomness and more complex cipher keys [2] because they may not resist too many conventional attacks such as brute force attacks. Hence, improving the randomness and keyspace requires brighter light of research. Also, a discrete chaotic system is often adopted as a catalyst for encryption schemes which lags in the complex behaviors, in which the generated sequences can be predicted easier.

Generally, the chaotic systems are classified as discrete and continuous chaotic systems. For all digital applications, continuous chaotic systems have to be converted into discrete data, and then, it should be digitized. There are many discretization methods like Euler method, Runge–Kutta method [3], etc. But, discrete chaotic systems are more attractive toward digital application. The behaviors of these systems slowly degenerate leading to the limited accuracy. In contrast, the continuous system has a complex structure, and this complex variable chaotic system can be applied in an encryption algorithm to improvise the security performance. Also, it can expand the variable space of the system that improves the dynamical characteristics. The complex chaotic system creates a challenging domain for malicious encoding. Hence, this system can be considered to enhance security performance against different malicious attacks [4, 5].

The classical chaos systems which hang about universally differ from the quantum chaos systems that rely on quantum mechanics theory. Quantum logistic system is built from the classical logistic system. Also, this logistic map uses the repetition procedure resulting in a chaotic behavior depending on its initial parameter. This system results in high dimension and complex dynamic behavior. Hence, the quantum logistic system is suitable for an efficient encryption algorithm. Enormous research papers have been proposed based on the quantum logistic system [6].

Considering all the above analysis, the novel encryption scheme Q-CNEST has been proposed which is based on a quantum chaotic system whose initial conditions are designed based on the dynamic varying network characteristics such as signal strength and power consumption of the system followed by the double tier permutation [7]. The following steps were adopted for the formation of novel encryption schemes.

1. Dynamic quantum chaotic key formation using network characteristics such as signal strength and power consumption modes.

2. Formation of new encrypted data with permutation between the cipher keys and information. (Tier-I mechanism)
3. Formation of new encrypted data by the process of diffusion (Tier-II mechanism).

The organization of the paper is as follows.

Section 2 deals with related works. Section 3 deals with the background work of the logistic maps. Section 4 describes the proposed architecture. Experimentation setup of the proposed system is explained in the Sect. 5, and the results with comparative analysis are discussed in Sect. 6. Conclusion along with future scope has been discussed in Sect. 7.

2 Related Work

Yi He et al. proposed a 2D spatiotemporal chaotic-based encryption algorithm. Both linear and nonlinear chaotic map lattices are considered. 2D-coupled map lattices and permutation process are incorporated to enhance security [8].

Moatsum et al. proposed a hybrid chaotic system using the perturbation process. In this system, confusion and diffusion processes are implemented involving pixel shuffling and substitution. Security and performance analysis are observed using various analyses [9].

Qiang Lai et al. proposed a 4D chaotic system. An infinite number of chaotic attractors is produced. The initial conditions for a chaotic system are obtained using a sine function. The attractors depend on the equilibria to determine the attraction location. They have suggested that increasing the equilibrium points may generate multiple attractors [10].

Ranjeet Kumar Singh et al. presented a novel method to utilize a security system for the transmission of computerized substance over the open system. The proposed strategy utilizes a staggered picture encryption/unscrambling calculation dependent on quantum confusion map and meager inspecting. In the underlying stage, a unique picture is partitioned into squares of equivalent estimate, and each square is additionally isolated into sub-square utilizing DWT method. The pixels of the nearby sub-squares are traded arbitrarily by an irregular network. Every recurrence band of the squares is encoded by modulus capacity and consolidated each to get the new square. Next, pixels of the neighbor squares are traded haphazardly by arbitrary network, and along these lines, each square is scrambled utilizing modulus work. The outcome given in graphical and unthinkable structure shows the properness of the calculation [11].

Guodong Ye et al. proposed a picture encryption calculation dependent on a confused guide and data entropy. In contrast to Fridrich’s structure, the proposed technique contains the change, regulation and dissemination tasks. This strategy stays away from the weakness in conventional plans of carefully rearranging the pixel positions before dispersion encryption. Data entropy is utilized to impact the age of the keystream. The underlying keys utilized in the change and dispersion stages communicate with one another. Thus, the calculation goes about as a resolute substance to upgrade security. Test results and security investigations exhibit the great execution of the proposed calculation as a safe and viable specialized strategy for pictures [12].

3 Background Work

3.1 Logistic Maps

Chaotic systems are characterized by initial sensitivity, randomness and high unpredictability. The mathematical expression for the logistic chaotic maps is given by

$$Xn + 1 = \mu Xn(1 - Xn) \tag{1}$$

where $0 \leq \mu \leq 4$ represents a bifurcation parameter. Generally, bifurcation occurs when a small change leads to an unpredictable system. A variation in a parameter leads to a change in differential system. An equilibrium may be unstable and results in either periodic solution or new stable equilibrium that makes the previous equilibrium unstable. The parameter which creates this change of state is called a bifurcation parameter. The initial value $x_0 \in (0,1)$ iterates the sequence $x1$, $x2$, ..., xn, when $3.5699456 < \mu \leq 4$ leads the logistic system *i* sin a chaotic state. Moreover, quantum logistic chaos which was proposed by Goggin et al. 1990 has the following mathematical expressions

$$\left\{ \begin{array}{c} a_{n+1} = C * (a_n - |a_n|) - C * f_n \\ b_{n+1} = -f_n * h^{-2\alpha} + h^{-\alpha} C[(2 - 2a_n) f_n - 2a_n d_n] \\ e_{n+1} = -d_n * h^{-2\alpha} + h^{-2\alpha} C[2(1 - a_n) d_n - 2a_n f_n - a_n] \end{array} \right\} \tag{2}$$

where C is the control parameter, α is the dissipation constant, a_n and d_n are considered to be the complex conjugate part of their real counterparts. By varying the initial conditions, the characteristics of the above-mentioned logistics chaotic maps also vary. When compared with the chaotic maps, quantum maps add the disturbance at the end of the process. Because of the sensitivity of the initial conditions, a very small change in information may lead to the production of different sequences. Due to the variation in each iteration, the above chaotic logistic maps will produce more nonlinear characteristics which suit the best for applications related to image encryption [13, 14].

4 Proposed Encryption Process

The whole quantum encryption process for the proposed scheme is as follows,

(a) Generation of pseudo-random generators using network-centric initial conditions in quantum logistic chaotic maps.
(b) Permutation process between the quantum chaotic maps and image information.
(c) Chaotic diffusion process.

4.1 Network-Centric Logistic Chaotic Maps

Since the paper proposes the encryption process for the next-generation networks such as the Internet of things (IoT) and long-term evaluation (LTE), the pseudo-random keys are generated by using the network-centric MAC parameters such as received signal strength (RSSI) and ID of the channel. The mathematical expressions for calculating the RSSI are given as follows

$$\text{Received Signal Strength Indicator (RSSI)} = -(10.\eta.\log(\mathrm{d}) + \mathrm{A}) \qquad (3)$$

where

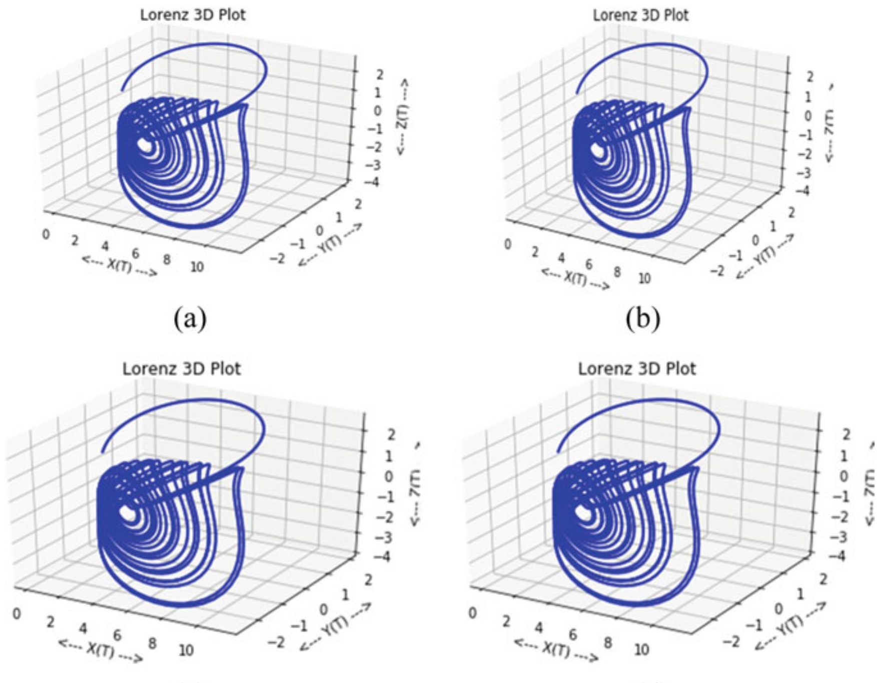
$$d = 10\left[\frac{(P_o - F_m - P_r - 10n\log(f) + 30n - 32.44)}{10n}\right] \qquad (4)$$

P_o Power of the signal (dBm) in the zero distance,
P_r Signal power (dBm) in the distance d,
F Signal frequency in MHz,
F_m Fade margin,
N Path-loss exponent.

These are the initial cmaps given in Eq. (2). Figure 1 represents different values of RSSI which are used as the initial conditions for chaotic behavior. Figure 1 illustrates the chaotic behavior of the proposed systems with different network parameters. From Fig. 1, it is clear that the system has unrelated chaotic circular behaviors with the different parameters which find its suitability in the encryption process.

4.2 Quantum Permutations

To implement strong protection, quantum-based XOR operation has been done. It has been implemented with the help of Hadamard basis and CNOT gate in QISKIT tool. The sequences which are obtained from the above process are sorted in ascending order without changing the nature of the bits [15, 16]. Let S = {S1, S2, S3, … Sn}



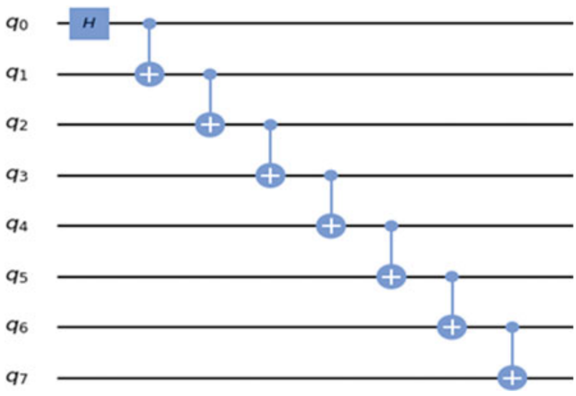
(a) (b)

(c) (d)

Fig. 1 Chaotic behaviors **a** 48 dbm, **b** 38 dbm, **c** 24 dbm, **d** 19 dbm

be the pseudo-random sequences and sorted sequences are given by S″ = {S1, S2, S3, S4, … Sn}. These sequences are then XORed with bit image information whose quantum implemented circuit is given in Fig. 2.

Fig. 2 Quantum circuit implementation of the permutated XOR bits for the secured transmission



4.3 Chaotıc Diffusion

To ensure the security process of the data, the diffusion process is employed between the new encrypted key from the above process and the input data streams [17, 18]. Before the diffusion process, the scaling process has to be done to satisfy the input data transmission time, such that all the output matrix elements must be scaled to 0–255 using Eq. (5).

The new dynamic constant β and diffusion operatorμ have been introduced in the diffusion process and input date streams, respectively, Eqs. (5)and (6).

$$\beta = \text{mod}\ 256 \left\{ \sum D(i) \right\} \quad \text{where} \quad i = 0, 1, 2, 3, \ldots 25 \tag{5}$$

$$\mu = Fi + \beta + \text{mod}\ 256\{\ d(i)\} \quad \text{where} \quad 0, 1, 2, 3, \ldots M \tag{6}$$

Hence, the new cipher data has been obtained through the diffusion process between the dynamic keys and the data streams.

4.4 Complete Encryption Process

Step 1: Medical images as input data streams

Step 2: Divide the input data streams into the length eight bytes which are scaled in accordance with the application (Our Case is 256).

Step 3: Formation of quantum logistics maps

Step 4: Measurement of MAC parameters such as RSSI, distance and channel ID.

Step 5: Formation of 'S' vector which has been formulated by quantum chaotic maps which depends on network-centric initial conditions.

Step 6: Rearranging the S-vectors in ascending orders.

Step 7: Formation of a new complex key by XORing the obtained the 'S' vector along with the image datasets.

Step 8: Rearranging the data as data matrix $d(i)$ and G as F matrix which are scaled to 256.

Step 9: Diffusion process is performed.

Step 10: Encryption process is completed.

5 Experimental Setup

For evaluating the proposed encryption scheme, the MIAS mammogram datasets have been used as the input images which consist of 322 mammogram images which are stored as PNG format. All the proposed algorithms are simulated in QISKIT

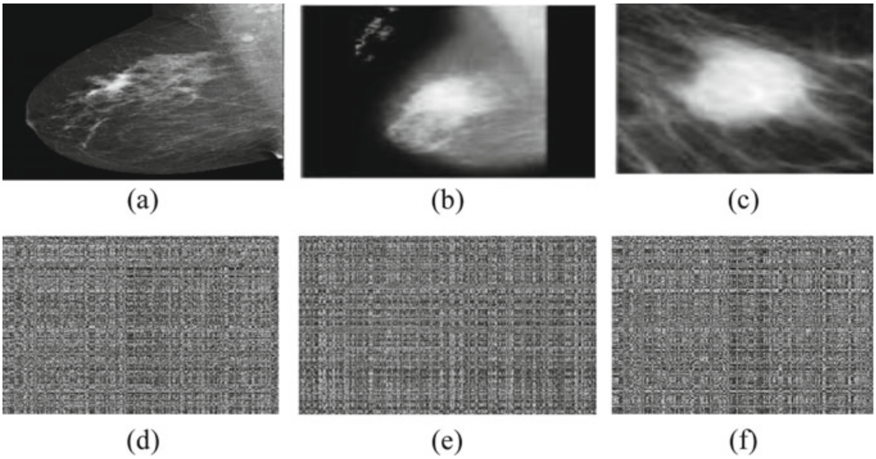


Fig. 3 **a** Normal Mammogram medical image, **b** Benign image, **c** Malignant image, **d** Encrypted normal image, **e** Encrypted benign image, **f** Encrypted malignant image

packages which run on Python 3.6 environment. The whole experimentation is implemented in computer which runs I7 CPU, 16 GB RAM, 2 TB HDD and 2 GB AMD Radeon GPU. Figure 3 shows different images used as the input for the proposed encryption scheme.

6 Various Analysis

6.1 Key Sensitivity Analysis

In this methodology, the secret key should be very sensitive even for minute changes resisting brute-force attack. To test the key sensitivity, the change in the number of bits has been observed with the change in the initial value of a chaotic system. The sensitivity of the key is calculated by the measurement of NPCR and UACI which measures the performance of this algorithm against the differential attack [19–21]. For a plain image, the mean value of NCPR and UACI can be calculated through an iterative procedure for nearly 20 times in a random manner. The mathematical expression for calculating the NPCR and UACI is given by the following equation. The analysis results are tabulated and discussed in Tables 1, 2 and 3.

$$\mathrm{NPCR} = \frac{\sum_{i,j} E(i, j)}{L} * 100 \tag{7}$$

Table 1 Illustration of NPCR and UACI for MAIS normal image datasets

S. No	No of bits change (%)	NPCR (%)	UACI (%)
1	10	99.4	89.90
2	20	99.3	85.57
3	30	99.2	86.90
4	40	99.1	85.56
5	50	99.0	85.89
6	60	99.1	86.90
7	70	99.2	89.90
8	80	99.1	89.56
9	90	99.2	88.90
10	100	99.1	88.34

Table 2 Illustration of NPCR and UACI for MAIS benign image datasets

S. No	No of bits change (%)	NPCR (%)	UACI (%)
1	10	99.34	88.80
2	20	99.25	85.47
3	30	99.2	86.00
4	40	99.10	85.06
5	50	99.10	84.79
6	60	99.23	83.00
7	70	99.28	87.00
8	80	99.19	88.00
9	90	99.10	84.90
10	100	99.45	83.34

Table 3 Illustration of NPCR and UACI for MAIS malignant image datasets

S. No	No of bits change (%)	NPCR (%)	UACI (%)
1	10	99.30	87.90
2	20	99.56	89.57
3	30	99.40	88.90
4	40	99.34	81.56
5	50	99.25	84.89
6	60	99.10	85.90
7	70	99.12	86.90
8	80	99.5	89.56
9	90	99.2	88.90
10	100	99.0	88.34

$$\mathrm{UACI} = \frac{1}{L}\sum_{i,j}\frac{|f(i,j) \neq f(i,j)|}{256} * 100 \tag{8}$$

where

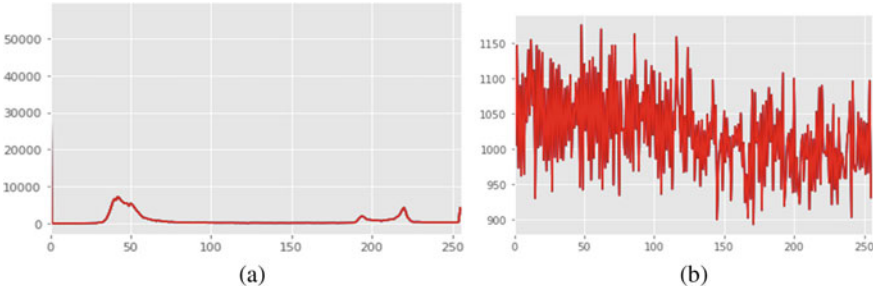
$$E(i,j) = \begin{bmatrix} 1, \ f(i,j) \neq f(i,j) \\ 0, \ f(i,j) = f(i,j) \end{bmatrix} \tag{9}$$

From the above tables, it is clear that the NPCR and UACI image datasets can resist different resistant attacks effectively. From the above tables, it is clear that the proposed chaotic encryption has good NPCR ranges from 99.56 to 99.0% even though the bit values are changed at different proportions. Entropy remains to be as low even for the different bit changes.

6.2 Statistical Performance Analysis

In this section, the statistical performance of the encryption algorithm has been observed to measure the degree of confidentiality of the image datasets. The histogram chart reflects the distribution of pixel values in the image. The pixels in closer result in better encryption. It illustrates the statistical figures in the image shown in Fig. 4.

Figure 4a represents the image before encryption which shows the uniformity of image used is very less in almost all parts of areas. Figure 4b represents the image after encryption which shows the uniform distribution of the images. The chi-square test is a method to measure the uniformity. The cipher images' distribution is calculated with critical thresholds with 25,15,10, 5 and 1% probability of changes whose values are given as 235.378, 244.29 and 249.44. Table 4 shows the chi-square value of different plaintext images and cipher images. From Table 4, it is clear that the



(a) (b)

Fig. 4 Histogram analysis results **a** image before encryption, **b** image after encryption

Table 4 Chi-square (α) distribution for the different image datasets

Image_details	Plain image	Cipher-images	Critical values				
			$\alpha_{(0.1)}$	$\alpha_{(0.5)}$	$\alpha_{(0.01)}$	$\alpha_{(0.02)}$	$\alpha_{(0.05)}$
Normal medical images	1.3456×10^5	235.378	Pass	Pass	Pass	Pass	Pass
Benign images	3.456×10^5	233.89	Pass	Pass	Pass	Pass	Pass
Malignant images	3.567×10^5	244.89	Pass	Pass	Pass	Pass	Pass

encryption has passed the above thresholds which consequently results in resistance of the statistical brutal force attack.

6.3 Adjacent Pixel Point Correlation Analysis

The adjacent pixel point correlation of the image can be analyzed by the following mathematical expressions.

$$R_{xy} = \frac{\operatorname{cov}(a, b)}{\sqrt{E(x)E(y)}} \tag{10}$$

$$\operatorname{cov}(a, b) = D\{[a - D(a)][b - D(b)]\} \tag{11}$$

$$e(a) = \frac{1}{n}\sum_{i=1}^{n} a_i \tag{12}$$

$$L(x) = \frac{1}{n}\sum_{i=1}^{n} [a_i - a(x)]^2 \tag{13}$$

where $e(a)$ and $L(x)$ represent the expectations and variance of the plain text images and cipher image datasets. The correlation between the plain images and cipher images are given in Table 5.

Table 5 Representation of the correlation coefficient analysis between the plain image and cipher images

Image_details	Plain images			Cipher images		
	Horizontal	Vertical	Diagonal	Horizontal	Vertical	Diagonal
Normal image	99.567	99.456	99.900	0.2333	1.6779	2.9000
Benign images	90.788	90.444	90.23	0.03445	0.56778	0.7889
Malignant images	89.890	89.00	85.56	0.00900	1.456	3.9090

From the above Table 5, it is clear that correlations between different views of the plain images remain to be same in the most viewing angles. But to contrast, encrypted images do not correlate with the other parts of the images. The difference between the correlation coefficients of the different views of normal images is found to be 0.001, but in case of encrypted images, it is found to be 1.5. Also, the difference of coefficient of malignant and benign images is found to be 1.4,10, whereas coefficient of encrypted images is greater than 1.4. From the table, it clear that correlation is high as 1.5 which increases the complexity in the location of the encrypted images to change.

6.4 Information Entropy Analysis

Data entropy is the measure of uncertainty which reflects the highest degree of uncertainty of image information. The higher values of entropy prove higher randomness of cipher images. The mathematical expression for the entropy calculation is given by

$$g(m) = \sum_{l}^{l-1} q(m) \log_2 \frac{1}{q(m_i)} \tag{14}$$

where l represents the gray level. $q(m) \rightarrow$ the probability of gray value that appears in the image matrix. As per theoretical concepts, for a 8-bit gray image, it is good to have an entropy value greater than or equivalent to 9 that shows the rate of unpredictability. Table 6 illustrates the value of information entropy for a different image.

The test results of entropy are listed in the above Table 6. For efficient data encryption, the value of entropy should be close to 10, and from the above table, the value of entropy is found to be 9.90 for normal images, 9.789 for benign and 9.78 56 malignant images. Based on the results obtained, the encrypted images cannot be decoded by an attacker.

Table 6 Illustration of different information entropies for image sets used

Image_datasets	Entropy	Local entropy (block of images)
Normal image sets	9.900	9.678
Benign	9.789	9.402
Malignant images	9.7856	9.211

Table 7 Computational complexity analysis

Image datasets	Computational time complexity
Normal image datasets	0.06 s
Benign image datasets	
Malignant image sets	

6.5 Computational Complexity Analysis

The time consumption for the proposed algorithm depends on two processes, namely permutation and diffusion process. The total time consumption for different image sets is depicted in Table 7.

The time of computation is calculated by the Python timers used for experimentation, and it is found to be 0.06 s which is considered to be faster operation and proves to be utilized in medical image transmission.

7 Conclusion and Future Scope

A novel network parameter-based quantum encryption process QCNET has been suggested for secured medical image transmission. The proposed encryption algorithm utilizes the permutation process to shuffle the image pixels. Furthermore, logistic mapped-chaotic diffusion is also accomplished to protect the images. The complex control factors of the logistic map help to develop a strong keyspace to resist against brute-force attack. The quantum circuit is designed, and numerical simulation results show that the proposed scheme is used to secure the information and resist various attacks. Moreover, the computational difficulty is lesser than other traditional encryption schemes.

The proposed algorithm has been tested only for one attack in an IoT environment. Still, the algorithm needs intelligence for different categories of attacks. To increase the encryption efficiency for various attacks, hybrid chaotic encryptions for multiple images can be researched further.

References

1. Liu X, Xiao D, Xiang Y (2019) Quantum image encryption using intra and inter bit permutation based on logistic map. IEEE Access 7:6937–6946. https://doi.org/10.1109/ACCESS.2018.2889896
2. Xu J, Li P, Yang F, Yan H (2019) High intensity image encryption scheme based on quantum logistic chaotic map and complex hyperchaotic system. IEEE Access 7:167904–167918
3. Zhang J, Huo D (2019) Image encryption algorithm based on quantum chaotic map and DNA coding. Multimedia Tools Appl 78:15605–15621

4. Hua Z, Zhou B, Zhou Y (2018) Sine-transform-based chaotic system with FPGA implementation. IEEE Trans Ind Electron 65(3):2557–2566 (March)
5. Axenides M, Floratos E, Nicolis S (2018) The quantum cat map on the modular discretization of extremal black hole horizons. Eur Phys J C 78(5):412–427
6. Zhang X, Seo S-H, Wang C (2018) A lightweight encryption method for privacy protection in surveillance videos. IEEE Access 6:18074–18087
7. Lia C, Zhanga Y, Xie EY (2019) When an attacker meets a cipher-image in 2018: a year in review. arXiv:1903.11764v2 [cs.CR] 5 June 2019
8. He Y, Zhang Y-Q, Wang X-Y (2020) A new image encryption algorithm based on two-dimensional spatiotemporal chaotic system. Neural Comput Appl 32:247–260
9. Alawidaa M, Samsudina A, Teha JS, Alkhwaldehb RS (2019) A new hybrid digital chaotic system with applications in image encryption. Sig Process 160:45–58 (July)
10. Lai Q, Chen C, Zhao X-W, Kengne J, Volos C (2019) Constructing chaotic system with multiple coexisting attractors. IEEE Access 7:24051–24056 (March)
11. Singh RK, Kumar B, Shaw DK, Khan DA (2018) Level by level image compression encryption algorithm based on Quantum chaos map. J King Saud Univ Comput Inf Sci
12. Ye G (2018) A chaotic image encryption algorithm based on information entropy. Int J Bifurcat Chaos 28(1):1850010 (11 pages)
13. Wang X, Pham V-T, Jafari S, Volos C, Munoz-Pacheco JM, Tlelo-Cuautle E (2017) A new chaotic system with stable equilibrium: from theoretical model to circuit implementation. IEEE Access 5:8851–8858 (June)
14. Aqeelurrehman X, Kulsoom LA, Ullah S (2016) A modified (Dual) fusion technique for image encryption using SHA-256 hash and multiple chaotic maps. Multimedia Tools Appl 75(18):11241–11266
15. Bakhshandeh A, Eslami Z (2013) An authenticated image encryption scheme based on chaotic maps and memory cellular automata. Optics Lasers Eng 51(6):665–673
16. Akhshani A, Akhavan A, Lim S, Hassan Z (2012) An image encryption scheme based on quantum logistic map. Commun Nonlinear Sci Numer Simul 17(12):4653–4661 (December)
17. Belazi A, El-Latif AAA, Belghith S (2016) A novel image encryption scheme based on substitution-permutation network and chaos. Sig Process 128:155–170
18. Kulsoom A, Xiao D, Aqeelurrehman, Abbas SA (2016) An efficient and noise resistive selective image encryption scheme for gray images based on chaotic maps and DNA complementary rules. Multimedia Tools Appl 75(1):1–23
19. Ji X, Bai S, Guo Y, Guo H (2015) A new security solution to JPEG using hyper-chaotic system and modified zigzag scan coding. Commun Nonlinear Sci Numer Simul 22(1):321–333
20. Wang XY, Zhang Y-Q, Zhao Y-Y (2015) A novel image encryption scheme based on 2-D logistic map and DNA sequence operations. Nonlinear Dyn 82(3):1269–1280
21. Ahmed F, Anees A, Abbas V, Siyal M (2014) A noisy channel tolerant image encryption scheme. Wirel Pers Commun 77:2771–2791

Secure Voting for Democratic Elections: A Blockchain-Based Approach

Hardik Ruparel, Shraddha Hosatti, Mahesh Shirole, and Sunil Bhirud

Abstract Blockchain technology, due to its lucrative features like immutability, transparency, and security, is definitely at the forefront of an impending digital revolution. Numerous organizations have already started integrating blockchain with their applications to enhance security. Currently, blockchain is being used in a wide spectrum of industries ranging from the energy domain to the automobile industry to the financial sector. Blockchain technology can be used in any domain that requires transparency, flexibility, and immutability as a fundamental requirement. One such domain targeted in this paper is the voting domain. Building a secure electronic voting system that offers transparency, immutability and security is the challenge that has been faced for a long time. This paper proposes a novel Blockchain-as-a-Voting-Service (BaaVS) solution that solves all the issues that are currently existing in the electronic-based voting system by delivering a secured and transparent voting approach.

Keywords Blockchain · E-voting system · Democratic elections · Secure immutable transparent vote · Flexible vote recounting · Ethereum

---

H. Ruparel (✉) · S. Hosatti · M. Shirole · S. Bhirud
Computer Science Engineering and Information Technology Department, Veermata Jijabai Technological Institute, Mumbai 400019, India
e-mail: hardikruparel14@gmail.com

S. Hosatti
e-mail: shraddha.hosatti@gmail.com

M. Shirole
e-mail: mrshirole@it.vjti.ac.in

S. Bhirud
e-mail: sgbhirud@ce.vjti.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_47

1 Introduction

Electronic votin g systems have been the subject of active research for decades, intending to minimize the cost of running an election, while ensuring the election integrity by fulfilling the security, privacy, and compliance requirements [6]. Although the evolution from pen paper-based voting to electronic voting systems has decreased the cost and enhanced the security manifold, there still exist some disadvantages like dependency on a centralized system, lack of transparency and no feature to ensure immutability, to name a few. Blockchain, with its inherent nature to provide transparency and immutability, perfectly fits as a solution to this problem. Blockchain technology has become the trending topic in the software world since the inception of Bitcoin [8] in 2008. A blockchain can be defined as a digital ledger of transactions that are duplicated and shared across the entire network of computer nodes. Conceptually speaking, it is a data structure that is built by a chain of blocks where each block contains a set of transactions. Along with the transaction data, cryptographic data like the hash of the previous block, timestamp of the transaction, Merkle root and nonce value is also stored in the blocks.

The process of adding a block to a blockchain is called *mining*. The mining of the blocks is typically a computationally expensive task that requires miner nodes to solve a cryptographic puzzle. The node that solves this puzzle first will propose the block to all the nodes in the network for verification. If the majority of the nodes verifies it positively and reach on a consensus, then the block will be added to the blockchain. It is the responsibility of the consensus algorithm to bring the blockchain network on to a decision.

Voting information needs to be unaltered and verifiable. Information stored in the blockchain is immutable and verifiable, hence blockchain is a good option for the voting system. Once a block is added to a blockchain, it cannot be amended retrospectively without the revision of all the subsequent blocks. This requires a consensus with a majority of the nodes in the blockchain network which is practically impossible [3], thereby making the blockchain immutable. Apart from being immutable, blockchain also provides the feature of easy verifiability. Since the blockchain ledger is distributed and duplicated across the nodes in the network, the nodes can easily verify the new block against their version of the ledger. These features are in part achieved through advanced cryptography, providing a security level greater than any previously known record-keeping system. Blockchain technology is therefore considered by many [1], including us, to have substantial potential as a tool for implementing a new modern voting process.

Although electronic elections are preferred for fast and secure operations, it is under threat from malicious actors that can infiltrate voting machines, alter voter registration databases, and more. A novel Blockchain-as-a-Voting-Service (BaaVS) architecture is proposed to leverage the benefits of the blockchain system in the voting system, thereby fabricating a new end to end secured voting system for democratic elections. The proposed system is fast, tamper-free, transparent and auditable. It is developed with the following objectives:

1. To store all stakeholders information in transparent and secure way
2. To enable secure anonymous voting with no double voting problem
3. To provide distributed, transparent, and re-countable voting Blockchain-as-a-Voting-Service (BaaVS) architecture.

The paper is organized as follows: Sect. 2 describes the related work in this domain. In Sect. 3 provides details about the proposed system. Section 4 presents the implementation details of the proposed system. In Sect. 5, the results of the implementation of the proposed system is presented. Finally, Sect. 6 concludes the paper.

2 Background and Motivation

Migrating the voting process online has been a challenging task for a long time. Many systems and architectures have been proposed to contest the voting process online. Each of these systems provides a different degree of privacy protection, transparency and risk mitigation. Various cryptographic techniques have been integrated to further fortify security. However, all the solution involves the presence of either a centralized system or a dependency of a third-party entity which makes the entire process vulnerable. Blockchain technology can be integrated into multiple areas such as cryptocurrencies, financial industry, games, anti-counterfeiting, supply chain management, etc. In blockchain-based secure voting system, a voter can utilize vote-token to cast his/her vote. There are myriad tokens [11], which are either fungible or non-fungible tokens are used to represent different types of assets in different industries. Some of the blockchain-based solutions that provide decentralized voting are as follows.

Open Vote Network [10] is a decentralized self-tallying voting protocol that is designed for a small-scale boardroom meeting. Unlike other voting protocols, Open Vote Network protocol does not rely on a third party for tallying. Each voter is in control of the privacy of their own vote. This voting protocol is a two-round protocol in which first the voters register their intent to vote in the elections post which they cast their vote in the second round. In [7], the open vote network protocol is implemented by developing smart contracts and deploying them on the Ethereum blockchain.

BitCongress [4] is another voting ecosystem that is created in conjugation with Bitcoin, Counterparty and Smart Contract blockchains using a distributed model to verify elections, votes and voters on separated blockchain networks. The smart contract blockchain is used to create elections as smart contracts that contain election protocols like election time, candidates, legislation and custom election rules. The counterparty blockchain is used to create addresses for the election smart contracts along with the addresses of all the actors of the system. It also keeps a tally of every vote cast in the election. Finally, there is the Bitcoin blockchain which is used for mining and storing the hash of the votes cast in the election. Tallying of the votes is done using the Borda count [4] mechanism.

In [3], a new blockchain-based voting system is proposed which provides security, privacy control measures and the ability to audit the votes. This system consists of three main modules viz. Identity Management, Cryptographic Privacy and Aggregation and Auditing. The identity management module is responsible for providing access control by requiring the voters to register first. The voter's registration details are stored on to a private blockchain which is then broadcasted to the election authorities. The cryptographic privacy module helps maintain the privacy of the voters by using the blind signature mechanism. The aggregation and auditing module comes into the picture after the election is over. All the valid votes are tallied, and the results are displayed. The voters will then be provided with the inverse function of the signature to verify and audit the votes.

A blockchain-based voting system has been proposed in [5] which allows the voter to change the vote during the election-time window even after the vote has been cast. The solution consists of a centralized system containing the list of all the qualified people who can vote. This system then distributes a security token to each of the qualified people which is then used as a digital signature. A vote that is cast is first encrypted using this digital signature and then is stored on to the blockchain.

Privacy Preserving Voting Protocol on Blockchain [13] consists of two main components viz. clients and smart contracts. The client component consists of the voting operations that will be performed by the voter, While the smart contract contains the voting logic and protocols. Each vote is first encrypted using hashing algorithms post which the validation process starts. In case of the vote has been proved invalid, the devoting steps are triggered. All the valid votes are then tallied and aggregated in the end.

After reviewing the related works in this domain, a new online voting system is proposed using blockchain to target the shortcomings of the current e-voting systems and provides benefits like increased security, transparency and ease of verifying the results.

3 Proposed Methodology

Our proposed system aims at bridging the shortcomings of the current online voting systems by providing a new Blockchain-as-a-Voting-Service (BaaVS) solution. Blockchain will ensure that once a transaction has been made, which is a vote, in this case, cannot be tampered with. Proposed Blockchain-as-a-Voting-Service (BaaVS) System's block diagram is shown in Fig. 1. The proposed voting system consists of six actors: voters, candidates, political party and Electoral Registration Officer (ERO), Polling Officer (PO) and Election Commission (EC).

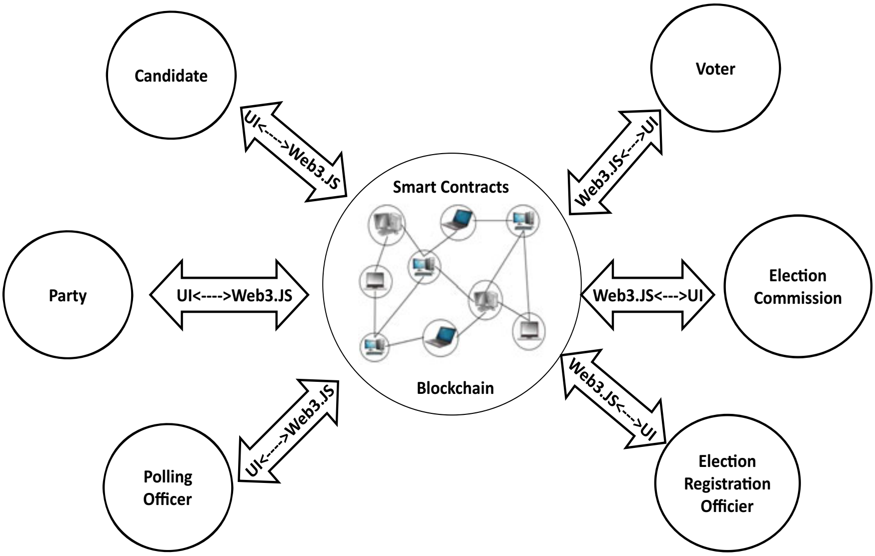


Fig. 1 Proposed blockchain-as-a-voting-service (BaaVS) system—Block diagram

3.1 Roles and Responsibilities

Voters, candidates and political parties have to register themselves by filling up the required information which is then validated by the contract deployed for registration. The ERO will then verify the details of the voters, the candidates and the parties. After a successful verification, the details entered by the actors are made permanent on blockchain for further usage. A smart contract is deployed on blockchain to ensure that a voter can vote only once thereby eradicating the double counting problem. The smart contract will also verify that the candidate has met all the conditions necessary to contest in the elections. The PO's responsibility is to tally the votes using proper cryptographic mechanisms.

Voter—A person is classified to be a voter only if he/she is the citizen of the nation and is above the voting age. The role of the voter is to register him/herself by entering the necessary details on the blockchain. After the successful verification, a voter will receive his/her voting credentials. The voter will get a vote-token for the respective upcoming election through which the voter will cast his/her vote that will then be stored in the blockchain forever. The voter can cast his/her vote through his/her through registered mobile election wallet application or he/she can approach the polling booths nearby during elections.

Party—A party is an organization that announces its candidates before every election process. The party must first register itself on the blockchain. Party must also announce their representative candidate before the election process starts. Party can also dismiss any of its candidates if the candidate is not verified.

Electoral Registration Officer—The responsibility of an ERO is to verify the details of the voters, candidates and also the parties. After a successful verification, the ERO will provide these actors with their cryptographic signatures which will be used by the system for signing the votes cast by the voters.

Polling Officer The PO is responsible to verify voters on voting booths in case the voter prefers to vote at polling booths. The PO plays a crucial role in the election process during voting and at the time of vote counting. During the voting phase, the system signs the votes cast by the voter by the public key. During the counting phase, he unlocks the transparent secure vote for counting by providing his private key.

Candidate—A candidate contests the election. A candidate is eligible for standing in the election only if he/she is a verified voter. Candidate first needs to register him/herself on the blockchain, after that he/she needs to submit a small security deposit for contesting the election. Candidate must belong to a valid political party. Based on circumstances he/she may withdraw candidacy at any point in time before the election as per the schedule by the Election Commission (EC).

Election Commission— Election Commission (EC) is responsible for declaring the elections, formulating the rules and schedule of the registration, voting and vote counting phases of the election.

Once the users register themselves onto the blockchain, they will be verified and then the voter can be allowed to vote in the election process. As the entire voting ecosystem is hosted on blockchain, the overall election process can be viewed by all the actors of the system. To implement the proposed voting ecosystem, smart contracts are developed using Solidity.

3.2 Overall Working of the Proposed Voting System

The proposed Blockchain-as-a-Voting-Service (BaaVS) ecosystem consists of three main phases namely the Pre-election phase, the In-election phase, and the Post-election phase. Fig. 2 displays the overall working of the proposed blockchain solution.

Pre-election Phase The Pre-election phase further consists of four sub-phases namely Voter Registration, Validation by ERO, Candidate Nomination and Candidate Withdrawal. Each of these sub-phases is discussed in detail below:

- Voter registration is the first sub-phase of the pre-election phase. All newly qualified voters will first have to register themselves on to the blockchain ledger. For registering, a voter must enter its Voter ID along with details like name, address, phone number, date of birth, etc. To enhance the security of the system, the voters are required to register their biometrics.
- Verification of voter by ERO, the second sub-phase, consists of the verification of the voter's details and biometrics by the ERO. Only after a successful verification, the ERO will access the secured Government servers to provide the voter with its

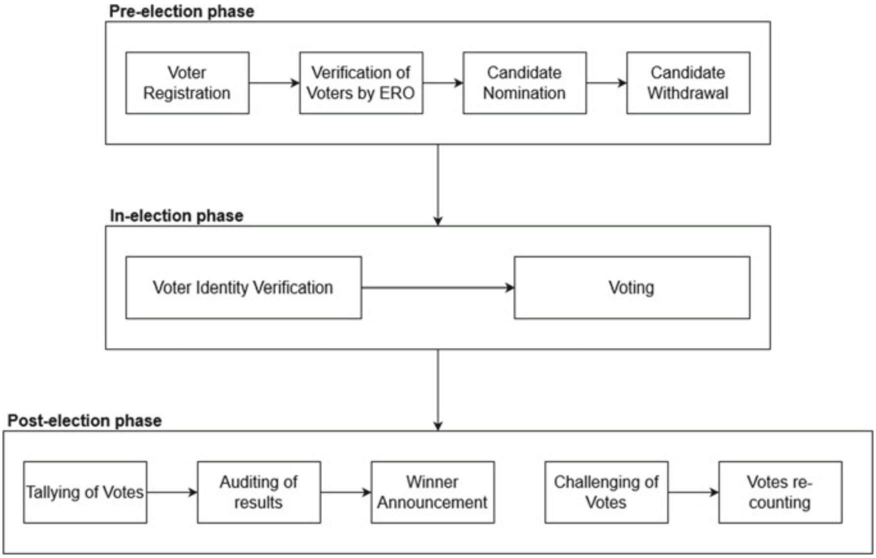


Fig. 2 Proposed system—Working

Private Key and its corresponding Public Key. Only the ERO will have access to the government's secured servers. The proposed system uses the private key of the voter to automatically sign the vote cast by the voter.

- Candidate nomination is the third sub-phase in the pre-election period during which all the political parties release a list of the candidates contesting the elections. Before nominating the candidates, the party must first register itself onto the blockchain. The candidates are also required to register themselves by entering the same details as that of the voter. Along with the above information, the candidate must also include the party he/she belongs to and a security deposit that must be submitted before contesting for any election. The candidate's deposit amount depends on the type of the elections he/she is contesting in.
- Candidate withdrawal is the final but an optional step in the pre-election phase. The party, at any point in time, can withdraw the candidacy of any candidate. This will lead to a refund of the security deposit collected during the candidate's nomination sub-phase.

In-election Phase After the Pre-election phase, all the voters, candidates and parties are registered and verified by the ERO. The in-election phase consists of Voter Identity Verification and Voting as two main sub-phases. The detailed functioning of each sub-phase is mentioned below:

- In the voter identity verification sub-phase, the ERO re-verifies the details entered by the voter during the Pre-election phase. After a successful verification, the ERO will verify if the voter belongs to the constituency he/she is intending to vote. For that, the ERO will verify the address mentioned on the voter's Aadhar Card. If the permanent address and the current address are same and the address belongs to the constituency the voter wants to cast vote in, then the ERO will show a green flag, thereby allowing the voter to vote. If the permanent address and the current address are not the same, but any one of these addresses belongs to the constituency, then too the ERO will allow the voter to proceed to the next sub-phase i.e. the voting sub-phase.

- Voting sub-phase lies at the core of the proposed Blockchain-as-a-Voting-Service (BaaVS). In this phase, the voter can cast his/her vote to any of its preferred candidates. Instead of capturing the votes using the traditional pen-paper approach or the electronic voting machine approach, the proposed system leverages blockchain to capture the votes. Before adding the vote in blockchain, the smart contract will first validate if this private key is used in any previous votes. If the system finds any other vote cast by the same private key, it will reject the vote. This validation allows the system to record only one vote from a particular voter, thereby solving the famous double-spending problem in the blockchain. To further fortify the security, the proposed system also signs the vote using the public key of the PO and at least three candidates contesting the elections as shown in Fig. 3. This Multi-signature authentication measure provides additional security and requires all the keys to tally the votes. Since the proposed voting system uses blockchain to store the votes, it allows voters to cast their vote online using any secured web-based cryptographic wallet. This is hugely beneficial for voters who cannot go to a voting booth to cast their votes. The proposed system uses voter's Aadhar number as a salt while calculating the cryptographic hash. This further strengthens the security by preventing the frequency analysis of the cryptographic hash.

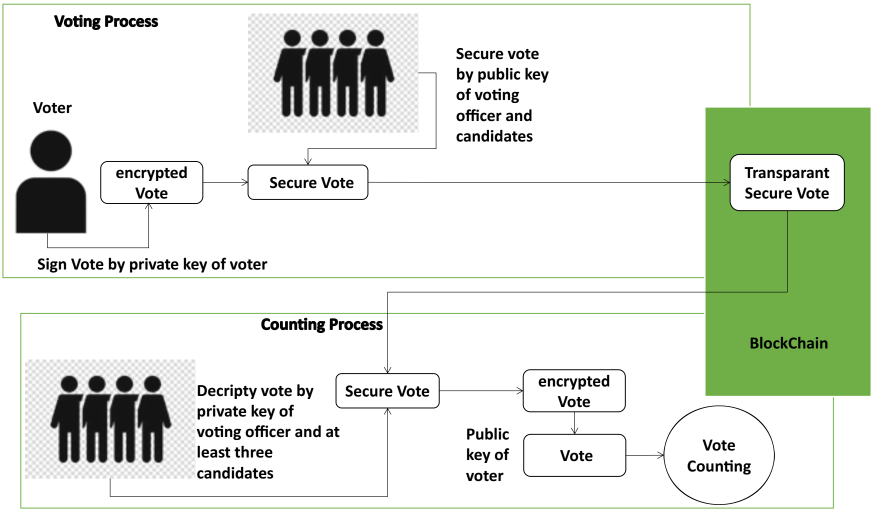

Voting Process
Voter
encrypted Vote
Secure Vote
Secure vote by public key of voting officer and candidates
Transparant Secure Vote
Sign Vote by private key of voter
BlockChain
Counting Process
Decripty vote by private key of voting officer and at least three candidates
Secure Vote
encrypted Vote
Public key of voter
Vote
Vote Counting

Fig. 3 A systematic view of secure voting and vote counting process

Post-election Phase This is the third and final phase of the proposed voting system. In a normal centralised voting ecosystem, the votes that are cast are tallied by a trusted third-party entity. This centralised dependency has a huge risk of the results getting manipulated since the votes cast by the voters are untraceable as the votes cast by the voters cannot be validated by the voters. To remove the third-party dependency for the vote tallying and to make the votes traceable and re-countable as many times as required, the benefits of blockchain are leveraged to solve these problems. A blockchain-based vote tallying mitigates the risk associated with the final vote count tampering. This phase consists of four sub-phases which are discussed in details below:

- Tally of Votes is the first sub-phase that the system performs after the election period. For the tallying process, the public key of the voter and the private keys of the PO and that of the same candidates are used to unlock the vote and tally it. This will allow removing the central dependency on just the PO, thereby providing a decentralized approach in tallying the votes.
- Auditing of the results provides the voters with an ability to themselves verify the correctness of their votes. Since, blockchain inherently provides features like immutability and transparency, the auditing of the votes can be easily be done by the voters on our proposed system.
- Winner announcement is the final sub-phase of the proposed system, in which the authorities announce the results.
- Challenging of votes is an optional sub-phase in the post-election period. This phase is only activated if anyone wishes to challenge the result and demands the recounting of the votes.
- Vote re-counting is an additional task performed by the BaaVS system to re-verify and validate the final results. Since all the votes are stored on the blockchain, the tallying of the votes can be done easily, cheaply and as many times as required which is a huge advantage over the current voting systems.

The proposed system, thus, solves the prevailing issues like security, tampering of the votes and transparency by using blockchain, which inherently provides all these benefits, to store the votes. The proposed system also provides additional security features like preventing double voting, multi-signature authentication and preventing the frequency analysis of the hashes which are extremely important to maintain the integrity of the elections.

4 Implementation Details

For implementing our proposed Blockchain-as-a-Voting-Service (BaaVS), smart contracts are developed. Smart contracts are developed using Solidity, a statically typed object-oriented programming language for developing smart contracts. These smart contracts are deployed on the ethereum blockchain [12] which are then executed by the miner nodes of the Ethereum network. During the in-election phase when the voter is casting its vote, along with the *candidateId* to which the voter wants to cast its vote to, the system ensures security by automatically signing the vote by the voter's private key, PO's public key and the public key of any two of the candidates. To avoid the frequency analyses of the hashes, the system also adds salt to the vote that needs to be hashed. In our case, the Aadhar Number of the voter are used as a salt. After digitally signing the vote and calculating the cryptographic hash, the hash is now ready to be mined on to the blockchain. The algorithm for the above mentioned steps of the voting sub-phase is given in Algorithm 1.

Since the system uses private/public keys of various actors to store the vote on the blockchain, during the vote tallying phase, the system will automatically use the corresponding public/private keys of the same actors to determine the *candidateId* to which the vote belongs to. Since cryptographic hashing is a one-way function, there is no way to find the *candidateId* to which the vote was cast just by using the cryptographic keys. The systems will loop through the entire list of *candidateId*

Algorithm 1 Casting Votes

```
function VOTE(candidateId, voterPrivateKey, POPublicKey,
candidate1PublicKey, candidate2PublicKey, salt)
Input: candidateId - ID of the candidate the voter has chosen to vote
voterPrivateKey- Voter's Private Key
POPublicKey - Public key of the PO
candidate1PublicKey- Public Key of any Candidate
candidate2PublicKey - Public Key of any other Candidate
salt- Salt for the Hash function (Aadhar number of the voter in our case)
Output: Cryptographic hash of the vote
    $hashOfVote \leftarrow$ Hash(candidateId, voterPrivateKey, POPublicKey,
candidate1PublicKey, candidate2PublicKey, salt)
    Add hashOfVote to the blockchain
end function
```

Algorithm 2 Algorithm for Tallying Votes

```
function TALLYING_VOTES(voterPublicKey, POPrivateKey, candidate1PrivateKey, candidate2PrivateKey, salt, hashOfVote)
Input: voterPublicKey - Voter's Public Key
POPrivateKey - Private Key of the PO
candidate1PrivateKey - Private key of any candidate
candidate2PrivateKey - Private Key of any other Candidate
salt - Salt for the Hash function (Aadhar number of the voter in our case)
hashOfVote - Cryptographic hash of the vote retrieved from blockchain
Output: Candidate ID for which the vote was cast
    for $candidateId$ in $candidateList$ do
        hash $\leftarrow$ Hash(candidateId, voterPublicKey, POPrivateKey,
candidate1PrivateKey, candidate2PrivateKey, salt)
        if ($hash = hashOfVote$) then
            return $candidateId$
        end if
    end for
end function
```

and perform cryptographic hashing functions on it. The count of the *candidateId* is increased by one if the hash matches the hash that is present on the blockchain. Algorithm 2 describes the procedure for tallying of votes that returns the *candidateId* for which the voter has cast the vote.

5 Results

For implementing the proposed system, five smart contracts- *Registration*, *Voting*, *Counting*, *Election_Configuration* and *Candidate_Nomination* are deployed on the Ethereum blockchain. As a result of the implementation, the performance of the system is measured by using a set of eight parameters which are described in Table 1.

5.1 Comparison with Existing Solution

The performance of the proposed system is compared with the performance of the current non-blockchain based solutions like Electronic Voting Machine (EVM), Electronic Polling Agent [3] and the proposed voting system in [2] against the parameters mentioned in Table 1. Due to blockchain's inherent nature of being immutable and transparent, the proposed system also inherits these features. Since the system is immutable, any vote that was cast cannot be changed at a later point in time, thereby maintaining the integrity of the elections. Hence, the proposed system, Electronic Polling Agent [3] as well as the proposed system in [2] provides integrity and transparency owing to the use of blockchain. Non-blockchain solutions like EVM cannot assure integrity and transparency features which are crucial for the election process.

The proposed system allows the user to verify whether his/her vote is being tallied correctly or not. The Electronic Polling [3] system also allows the voters to verify their votes by using the inverse function of the signatures. The blockchain-based

Table 1 Parameters for determining the performance of the system

Parameters	Definition
Integrity	Votes should not be modified after being cast by the voter
Transparency	Voters should have a general view of the entire process
Verifiability	Voters must be able to verify that the votes are correctly tallied
Cost-effectiveness	System should be affordable and efficient
Votes recounting	System's ability to recount the votes as many times as required
Scalability	System must be able to handle large number of transaction
Enhanced security	System's ability to provide advanced security features
Secrecy	No one should be able to determine how any individual has voted

voting system in [2] provides the voters with an ability to verify his/her votes by using cryptographic mechanisms. However, in non-blockchain solutions, there is no way for the voters to verify that their votes are correctly tallied or not. Leveraging blockchain has not only led to a significant decrease in the cost but also improved the transparency in the voting and the tallying phase. Hence all the blockchain-based solutions are cost-effective while the non-blockchain solutions are not. Votes recounting is a fundamental feature included in the proposed system which leverages the smart contract to allow efficient recounting of the votes as many times as necessary. The smart contract loops through the entire blockchain for getting the cryptographic hashes of the votes and uses Algorithm 2 to determine the *candidateId* for which the vote belongs to. Moreover, in the Electronic Polling Agent [3] and the blockchain-based system in [2], one cryptographic hash of the vote is stored in one block. This affects the scalability of the system as the miner nodes will only be able to process and mine one vote at a time. However, the proposed system allows several hundred cryptographic hashes of the vote to be processed concurrently by the miner nodes, thereby increasing the scalability of the system.

Additionally, the proposed solution provides advanced security features like preventing frequency analysis of the hash and multi-signature authentication. In the absence of salt in the hash function, a malicious user can look at the cryptographic hash of the votes and can identify the trend of the ongoing elections. The proposed system mitigates this risk by mandating the use of the salt while creating the cryptographic hash of the vote. Using the salt also helps maintain the true secrecy of the election as no one will be able to determine how the voter has voted. Thus, the proposed voting system provides advanced security features along with maintaining true secrecy by using additional cryptographic measures which are not present in Electronic Polling Agent [3] and the voting system in [2]. Non-blockchain solutions provide secrecy owing to the absence of transparency in it. Table 2 provides the comparative analysis and the benefits of the proposed system over some of the existing blockchain and non-blockchain based solutions.

However, due to blockchain's limited scalability, additional scalability measures need to be taken to allow the processing of millions of votes. For this, it is required to

Table 2 Comparative analysis with related systems [2, 3]

Parameters	Non-blockchain solutions	[2]	[3]	Proposed system
Integrity	×	✓	✓	✓
Transparency	×	✓	✓	✓
Verifiability	×	✓	✓	✓
Cost-effectiveness	×	✓	✓	✓
Votes recounting	✓	×	×	✓
Scalability	×	×	×	✓
Enhanced security	×	×	×	✓
Secrecy	✓	×	×	✓

implement advanced scalability measures like GeoSharding along with our proposed system. GeoSharding [9] is a sharding protocol that divides the blockchain network in shards and then elects a leader in each shard for parallel processing of the transaction which will help increase the throughput of the blockchain system thereby increasing the capacity of our system to enable millions of voters to vote concurrently.

6 Conclusion

This paper introduced a blockchain-based, secured electronic voting system that utilizes smart contracts to enable secure and cost-efficient election while guaranteeing voter's privacy. Due to blockchain's inherent feature of immutability, the data related to each election conducted using the proposed online voting system is safe, immutable, irreversible, and stored permanently in the blockchain ledger. The proposed solution provides additional advantages like multi-signature authentication, prevention of hash frequency analysis and double voting that is beneficial to all the stakeholders of the system. Leveraging blockchain's security features has offered a variety of new possibilities to overcome the limitations and adoption barriers of electronic voting systems, thereby ensuring election security and integrity and laying the ground for transparency. The proposed system allows the voters to validate their respective votes and candidates to view the tally process themselves. Moreover, using an Ethereum private blockchain has made it possible to send hundreds of transactions per second onto the blockchain, utilizing every aspect of the smart contract to ease the load on the blockchain. For countries of greater size like India, the election process spans out for multiple days for different geographical regions. One day, one election for huge democracies can be possible only if some advanced scalability measures are implemented along with our proposed system.

References

1. Weaver N Secure the vote today. https://www.lawfareblog.com/secure-vote-today
2. Ayed AB (2017) A conceptual secure blockchain-based electronic voting system. Int J Netw Secur Appl 9(3):01–09
3. Babu A, Dhore VD (2020) Electronic polling agent using blockchain: a new approach. In: IC-BCT 2019. Springer, pp 69–77
4. BitCongress: Bitcongress—Blockchain based voting system. http://cryptochainuni.com/wp-content/uploads/BitCongress-Whitepaper.pdf
5. Hardwick FS, Gioulis A, Akram RN, Markantonakis K (2018) E-voting with blockchain: an e-voting protocol with decentralisation and voter privacy. In: 2018 IEEE International conference on Internet of Things (IoT) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData). IEEE, pp 1561–1567
6. Hjálmarsson F, Hreiarsson GK, Hamdaqa M, Hjálmtýsson G (2018) Blockchain-based e-voting system. In: 2018 IEEE 11th international conference on cloud computing (CLOUD). IEEE, pp 983–986

7. McCorry P, Shahandashti SF, Hao F (2017) A smart contract for boardroom voting with maximum voter privacy. In: International conference on financial cryptography and data security, Springer, pp 357–375
8. Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system. https://bitcoin.org/bitcoin.pdf
9. Ruparel H, Chiplunkar S, Shah S, Goradia M, Shirole M (2020) Geosharding—A machine learning-based sharding protocol. In: IC-BCT 2019. Springer, pp 105–118
10. Seifelnasr M, Galal HS, Youssef AM (2020) Scalable open-vote network on ethereum. IACR Cryptol 2020:33
11. Shirole M, Darisi M, Bhirud S (2020) Cryptocurrency token: an overview. In: IC-BCT 2019. Springer, Singapore, Singapore, pp 133–140
12. Wood G et al (2014) Ethereum: A secure decentralised generalised transaction ledger. In: Ethereum project yellow paper, vol 151, issue 2014, pp 1–32
13. Zhang W, Yuan Y, Hu Y, Huang S, Cao S, Chopra A, Huang S (2018) A privacy-preserving voting protocol on blockchain. In: 2018 IEEE11th international conference on cloud computing (CLOUD). IEEE, pp 401–408

Convergence Analysis of Self-Adaptive Equalizers Using Evolutionary Programming (EP) and Least Mean Square (LMS)

N. Shwetha and Manoj Priyatham

Abstract Digital communication has become an important part of our lives, and technology has been undergoing advancements. With the arrival of the age of digitalization and digital signal, communication has got implemented in a vibrant range of applications but still, they are strongly affected by two basic problems, namely *Noise and Inter-Symbol Interference* (*ISI*). This is caused by the error-creating phenomena which are characteristics between the transmitter and receiver which include the scattering of the transmitted signal. The noise produced in the communication channel is caused by channel characteristics and can be reduced with proper channel selection. The SNR can be improved by improving the transmitter signal strength even in spite of noisy signal at the receiver. By using the adaptive equalization in channels will reduce this effect drastically and can be implemented by using various adaptive algorithms. Hence, an adaptive channel equalizer is used to inverse the effect channel had on the signal to get back the initial information. There are many adaptive algorithms to update the coefficients of equalizers; evolutionary algorithms are used in this paper to do so. The two algorithms used before are Artificial Bee Colony algorithm (ABC) and Ant Colony Optimization (ACO). The latest algorithm is the combination of Evolutionary Programming and LMS algorithm (EPLMS); this gives better solution faster. A comparative study between the algorithms is done in this paper.

Keywords Intersymbol interference (ISI) · Least mean square (LMS) · Artificial bee colony algorithm (ABC) · Ant colony optimization (ACO) · Evolutionary programming (EP) · Maximum likelihood sequence estimation (MLSE)

N. Shwetha (✉)
Department of ECE, Dr. Ambedkar Institute of Technology, Bangalore, Karnataka 560056, India
e-mail: shwethaec48@gmail.com

M. Priyatham
Department of ECE, APS College of Engineering, Bangalore, Karnataka 560082, India
e-mail: manojopriyatham2k4@yahoo.co.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_48

1 Introduction

Modern digital communication systems enforce the usage of channel equalization with a high tracking rate and short preparation time. Such constraints focus our interest on adaptive algorithms that are used to unite quickly. The most fundamental advantages of the digital system for video, information, and voice associations are their superior consistency in noise environment conversely with that of their analogy components; unfortunately, the most frequently digital communication of data is complemented with a trend commonly referred to as Inter-Symbol Interference (ISI) [1–3]. Momentarily, this implies that the transferred pulses are covered out so that the pulses that parallel to various signs are not distinguishable. Dependent on the broadcasting media, the most important reasons for ISI are wired communication; in reality, they are band restricted and has multichannel dissemination.

It is important to diminish the impacts of ISI for a dependable digital broadcasting system wherein the adaptive equalization comes into picture. Two of the more intensively developing fields of digital communication, specifically cellular communications and digital subscriber lines, are heavily reliant on the implementation of reliable channel equalizers. The LMS algorithm as seen is one of the extremely prevalent algorithms in adaptive signal handling. It was the emphasis of a lot of research and its realization in several applications because of its robustness and simplicity. One of the potential solutions is the enactment of the equalizer by a filter with finite impulse response (FIR) utilizing the perfectly established LMS algorithm for adapting its coefficients. The recognition stems from its comparatively low computer-based intricacy, good mathematical strength, simple configuration, and simplicity of implementation in accordance with the conditions of the devices. The principle of the LMS algorithm is to modernize the coefficients of adaptive filter recursively together with the adverse descent of the assessment of error surface. The traditional algorithm utilizes a static step size to accomplish the repetition and to find a deal among the divergence of small stable-state MSE and quick convergence. A little step size could make sure little MSE with a sluggish convergence, while a huge step size will offer improved tracking abilities and a faster convergence at the expense of greater stable-state MSE. Consequently, in flexible step-size LMS algorithm, it is impossible to resolve this ambiguity. Therefore, several variable step-size algorithms have been recommended to resolve the trouble [4–7]. Although such algorithms could speed up the convergence and determine stable-state MSE to a certain extent, they have been unsuccessful to examine the optimization of variable step-size LMS additionally. The following equations have been utilized to describe the LMS algorithm.

$$e(k) = d(k) - WT(k) * X(k) \tag{1}$$

$$X(k+1) = X(k) + W(k)\mu e(k) \tag{2}$$

where the dimension of $X(k)$ is the extension of the AF, $e(k)$ is deviation error, $d(k)$ is the anticipated output value, $X(k)$ is coefficient vector of the adaptive filter, $W(k)$ is the *i*/*p* vector at random sampling time n, and μ is learning step.

2 Concept of Inter-Symbol Interference

In the digital communication system, if everything is right at the receiver side then there will be no interaction among successive symbols. Here each of the signals which are arrived is decoded self-reliantly among others. But when it comes to symbol interaction, one of the waveforms will corrupt the values of the next nearby symbols. Due to this, the received signal will be distorted. Because of this, it is difficult to differentiate messages from such a received signal. The shortage is identified as the ISI. The purpose of an equalizer is to reduce the ISI so that a reconstructed signal having from the transmitter side. Due to this, it also reduces the bit rate of the transmitted signal. As assumption made in all pass AWGN is impractical, the lack of frequency spectrum the signal is filtered to minimize the bandwidth so that frequency structured division can be obtained [8, 9]. There are many bandpass channels available in practical but the response varies to the different frequency components. To avoid this, the simplest AWGN model is needed to have for representing the practical channels very accurately. Such commonly available retirement is a dispersive channel model shown in Fig. 1.

$$y(t) = x(t) * h_c(t) + n(t) \tag{3}$$

From the equation, $u(t)$ is the *Tx* signal, $h_c(t)$ is the impulse response of the channel & $n(t)$ is AWGN power spectral density. The dispersive representative of the channel is prototyped by using the linear filter $h_c(t)$. This dispersive channel model is a low-pass filter. By using this low-pass filter as can line the transmitted signal to time causing the effect of symbol difficult to adjust symbols in a practical case while transmitting the signals from the transmitter. Due to this, the ISI will deteriorate the error caused by the transmitted signal to error performance in the communication

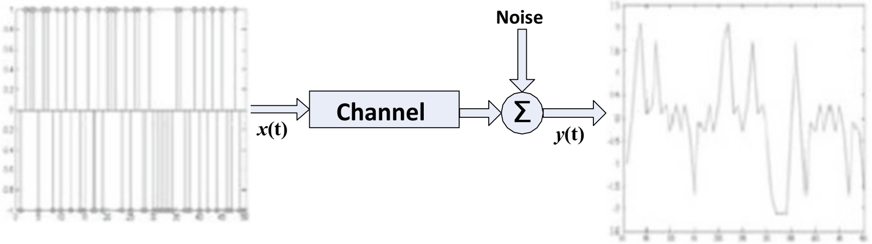


Fig. 1 Inter-Symbol Interference

system. Two main methods are mainly concentrated on which eradicates the ISI deterioration effect. In the first method, the band restricted transmission pulses are used to diminish the ISI. The pulses obtained by the ISI are called free pulsed which are known by its name Nyquist pulses [10–13]. As seen in the second method, the received signal is needed to screen to stop the ISI which was presented by the channel impulse response. This is known as equalization.

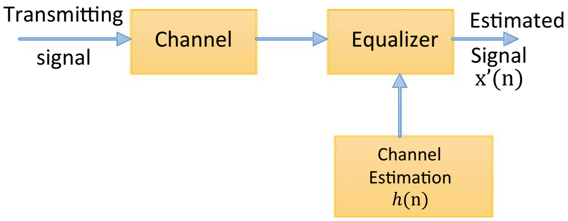
3 Process of Equalization

The equalization process is something like adjusting the balance between frequency components. The receiver signal matching frequency at the receiver is adjusted to decrease the noise and any interference generated during transmission. Hence, post-compensation of ISI is done at the receiver side by equalization. An equalizer is a device which does equalization. There are many applications of equalizers in electronics and communication. Here equalizing the equalizer to the channel will be the main concentration.

3.1 Channel Equalization

Channel equalization is one method of adjusting the equalizer coefficients to channel coefficients to reduce ISI. If a channel is considered as a filter, then equalizer is an inverse filter. An equalizer is not only going to compensate the effect of the channel but also going to compensate all the unnecessary effects transmitted signal went through, i.e., due to pulse shaping, transmitter filter, and receiver filter to get back the initial signal. The block diagram of channel equalization is shown in Fig. 2. When the channel is known, then by sending a known signal through the channel error signal can be calculated [14, 15]. Always the received signals differentiated to desire signal to receiver signal to identify the errors in the signal. The error signal is the driving force for the equalizer. Equalizer will aim to minimize the error signal. Hence, optimization techniques/algorithms are used to achieve this. There are many algorithms used for equalization. The most effective algorithm used before the adaptive algorithm is Maximum Likelihood Sequence Estimation (MLSE), where depending on

Fig. 2 Channel equalization



the MLSE algorithm the channel response for impulse is measured. The equalizer coefficients are adjusted or equalized to nullify the effect by channel. Adjustment of the coefficients is done to reduce the ISI and noise at the output. Hence, from the distorted version of the transmitted signal, the original version can be reconstructed by equalizer [16, 17]. Once the equalizer weights are set, then it won't change and then required information can be sent through the channel.

3.2 Adaptive Channel Equalizer

If the channel is time-variant, then the weights/coefficients used in the equalizer should also be updated as per the varying nature of the channel. Adaptive equalizer works in two modes: exercise mode (Training Mode) and choice directed mode (Decision Mode).

(a) Training mode: In this training, the signal is used, i.e., it is known to both transmitter and receiver. When equalizer gives an output, this is compared with the training signal and difference is the error signal. This error signal is used to update the equalizer weights/coefficients. When the error signal is zero, i.e., the output from the equalizer is equal to the training signal. Hence, the coefficients are saturated. Then, the equalizer changes its mode.

(b) Decision directed mode: Now the actual data signal can be transmitted from the transmitter. Again the equalizer changes its mode to training mode.

Again the weights will be updated and then in decision-directed mode, the next set of actual data will be sent. Therefore, the equalizer will be switching between these modes. The weights are varying with each new training mode. This forms a kind of feedback loop. This makes the communication system reliable. There are many adaptive algorithms which tell how to update the equalizer coefficients.

Figure 3 demonstrates the standard template for a direct channel with an adaptive equalizer. From the figure, it is observed that inverse filtering or channel equalizer comprises assessing a transmission functionality to counteract for the linear deformation triggered by the channel. From a different perspective, the goal is to compel a specified dynamic performance for the cascading of the adaptive filter and the channel

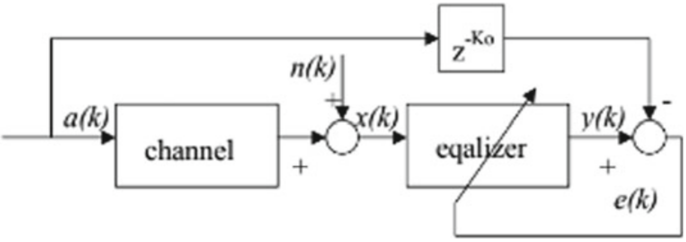


Fig. 3 Adaptive channel equalizer

(an unidentified system), decided by the incoming signal. The initial translation and interpreting are additionally relevant in communications, in which the data is sent via dispersion channels [18–20]. The next is suitable for management applications, in which the inverted filtering program creates control signals that will be utilized in the undetermined system.

Let us describe the vector of weighting coefficients $X(n)$ = $[X0(n)$ $X1(n)$ $X2(n)\ldots XM - 1(n)]T$ and the signal vector at equalizer input $W(n)$ = $[w(n)$ $w(n - 1)$ $w(n - 2)\ldots w(n - M + 1)]T$ of the adaptive filter at a moment n, [1–8]. Furthermore, at the equalizer the input signal samples are in the type of:

$$W(n) = W j h(j) a(n - j) + k(n) \tag{4}$$

where $h(j)$ signifies the response of channel impulse, $k(n)$ signifies the additive noise with the variance $\sigma 2k$, and $a(n)$ signifies the nth data sample.

The data specimens hold on values of ±1 ($a(n) = \pm 1$) only, and the noise is supposed to be autonomous. At the nth iteration instant, and the equalizer output is given as:

$$y(n) = XT(n)W(n) \tag{5}$$

The outcome $y(n)$ is being utilized in assessing the data transferred symbol $a(n - no)$, with N_0 signifying the interruption. The error sampling of kth output is:

$$e(n) = y(n) - a(n - N_0). \tag{6}$$

The coefficients of weighting in the LMS algorithm are derived from the subsequent statement:

$$X(n + 1) = X(n) + \mu e(n)W(n), \tag{7}$$

where μ is the algorithm step size. The output mean square error (MSE) is:

$$e(n) = E(e2(n)) = XT(n)RX(n) + E(a2(n)) \tag{8}$$

$$2XT(n)E\big(W^{-}(n)a(n - no)\big),$$

with $R = E(W^{-}(k)X^{-}T(n))$.

The regular output MSE after nth iteration may be stated as:

$$\varepsilon\text{avr}(n) = \varepsilon\text{MIN} + E\big(VT(n)RV^{-}(n)\big), \tag{9}$$

where $V(n) = X(n) - X{*}(n)$ is the error vector of the weighting coefficient and εMIN is the minimal MSE (5), for optimum vector for weighting coefficients is $X{*}(n)$, i.e., Wiener vector. The MSE beyond εMIN is referred to as the surplus MSE in the stable state. As demonstrated in [1, 2], for the LMS algorithm the surplus MSE is provided by:

$$\varepsilon \mathrm{e} \;= 12\mu\sigma 2\ \mathrm{ntr(R)}. \tag{10}$$

It can be noticed from (7) that the excess MSE is proportionate to the step size because of the gradient noise. The step size should be chosen to strike an equilibrium among the contradictory objectives and targets of the small stable-state error and fast convergence (large step size), i.e., little surplus MSE (small step size).

4 Problem Formulation

Realizing the period size is enormous, merging rate of the LMS algorithm will be depraved, yet the consistent state MSE, i.e., the mean square error resolves itself. Then, again will be more, if and only if the step size is small, the consistent state MSE will be small, yet the merging rate will be moderate. In this way, the step size gives a trade-off among the merging rate and the consistent state MSE of the LMS algorithm. The other way to increase the efficiency of the LMS algorithm is to make the step size adjustable as opposed to fixed which leads to VSSLMS algorithms. By using this methodology, both the fast merging rate and a little consistent state MSE can be achieved. The step size should satisfy the condition:

0 < step-size < 1/(max Eigenvalue of the input auto-correction matrix).

For fast convergence, step size is set close to its maximum allowed value.

5 Formulation of LMS Algorithm

The LMS, i.e., least mean squares procedure, is one of the utmost famous procedures in adaptive handling of the signal. Because of its robustness and minimalism was the focal point of many examinations, prompting its execution in numerous applications. LMS algorithm is a linear adaptive filtering algorithm that fundamentally comprises of two filtering procedure, which includes calculation of a transverse filter delivered by a lot of tap inputs and creating an estimation error by contrasting this output with an ideal reaction. The subsequent advance is an adaptive procedure, which includes the programmed modifications of the tap loads of the channel as per the estimation error. The LMS algorithm is additionally utilized for refreshing channel coefficients. The benefits of the LMS algorithm are minimal calculations on the sophisticated nature, wonderful statistical reliability, straightforward structure, and simplicity of usage regarding equipment. LMS algorithm is experiencing problems regarding step size to defeat that EP, i.e., evolutionary programming is utilized (Fig. 4).

Basic steps in the LMS algorithm:

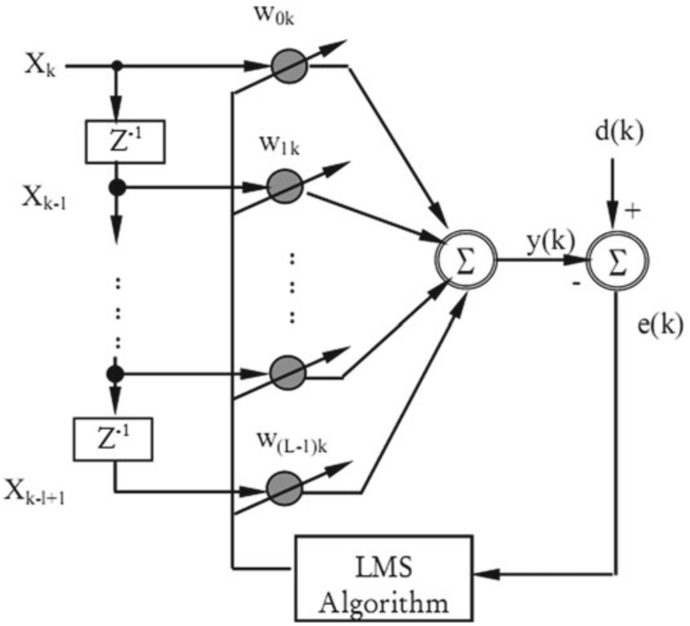


Fig. 4 Adaptive filter using LMS algorithm

1. Fundamentally at the beginning, random coefficients are taken.
2. Error is defined to the present sample using the equation

$$e(n) = d(n) - x^{\mathrm{T}}(n) * W(n)$$

where

$e(n)$	deviation error
$d(n)$	expected output value
$x(n)$	*i*/*p* vector at sampling time $nW(n)$ coefficient vector

3. Adjustment has defined using the LMS.
4. To get the new coefficient value.

$$W(n+1) = W(n) + \mu e(n)x(n)$$

5. New coefficients will replace the previous coefficients and the process will continue from step 2.

6 Evolutionary Programming

Evolutionary algorithms are stochastic search methods and not the deterministic ones. In 1960, Lawrence J. Fogel utilized the evolutionary programming in the USA to utilize modeled evolution as an educational procedure which is seeking to create AI. The previously existing methods like linear programming, calculus-based methods, for example, nutenian method are having difficulties in delivering the global solution. They are tending to stuck in the local solution. To overcome this problem, nature inspiration computation can be applied. In this approach, some characteristics that are available in nature are taken as a reference to develop the mathematical model. This mathematical model will utilize to find out the solution to the problem. In this paper, the characteristics of nature are taken as evolution [21, 22]. This is one of the most successful characteristics available in nature where the things evolved (the things changed) with the time to adapt the environment in result betterment in fitness value hence, the chances of survival are high, for example, the transformation from a monkey to a human. The mathematical model based on evolution is referred to as evolutionary computation (Fig. 5).

6.1 Basic Steps in Evolutionary Programming

1. Based upon natural evolution, a mathematical model called evolutionary computation has created.
2. In nature, the things change from one time to others to increase its fitness so that chances of survival could be better, for example, human evolution.
3. At the beginning, a random population is defined as a set of solutions.
4. Depends upon the objective a mathematical function called—objective function is defined.
5. With the certain mathematical operator called—mutation offspring's are generated.
6. A mathematical operator called selection operator is applied to define the next-generation members.

The simple and efficient way to achieve the better next generation is

a. Combine parent and offspring population.
b. Get their fitness.
c. Pick up the half members having higher fitness(tournament selection process is used).

7. The next generation will replace the previous generation to produce offspring.
8. This process will keep continuing until some terminating criteria do not satisfy.
9. After termination from the last generation, the member having the highest fitness will consider as a final solution.

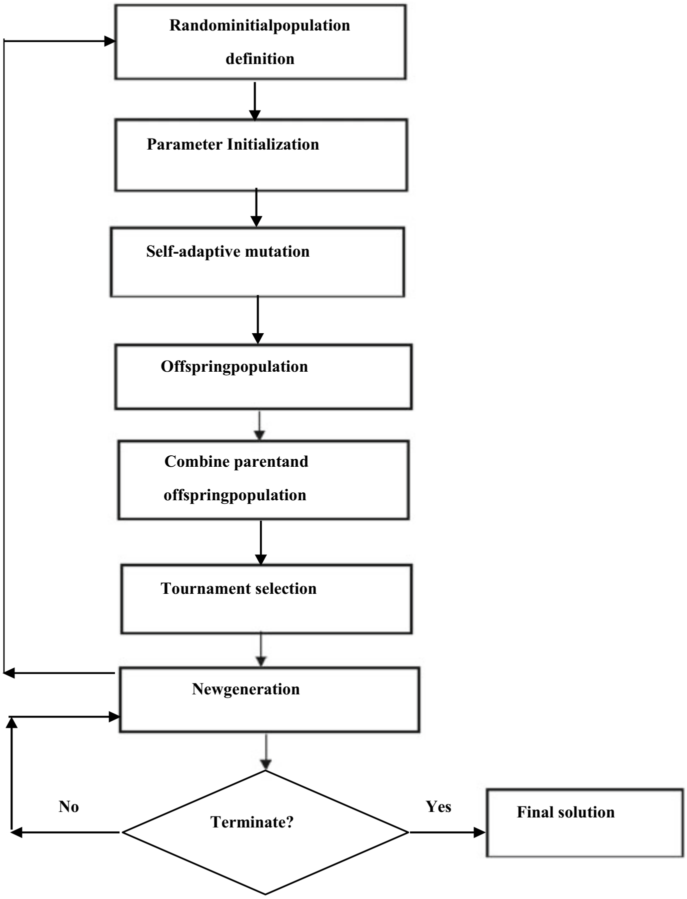


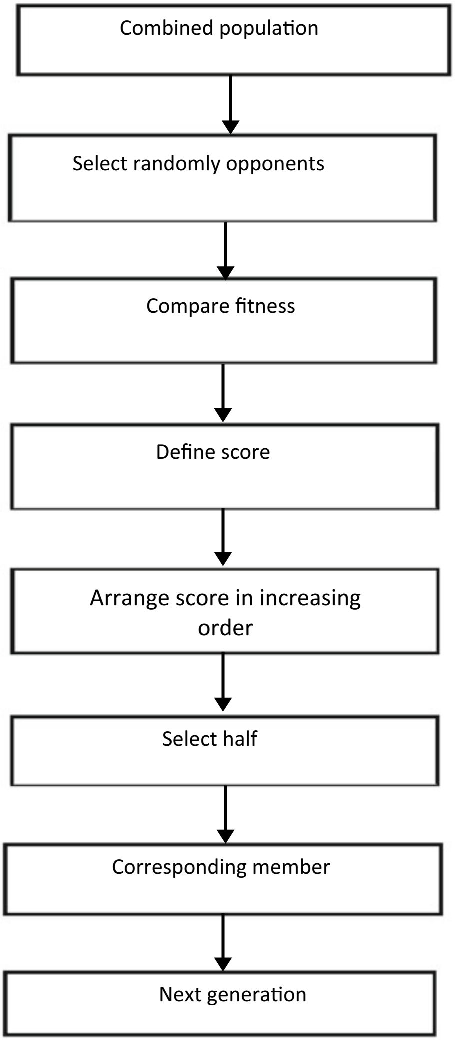
Fig. 5 Flowchart for evolutionary programming

6.2 Tournament Selection

6.2.1 Procedural Steps

1. Take the combined population of parents and offspring.
2. Pick up one member randomly and select ‗n‘no of opponents randomly.
3. Compare fitness.
4. Calculate the score.
5. Sort the score in the ascending order.
6. Select the right half.
7. Then pick up the members corresponding to that score.
8. These members will form the next generation.

Fig. 6 Flowchart for tournament selection



9. The next generation should replace the previous generation to produce offspring (Fig. 6).

7 Working of EP-Based LMS Algorithms

1. At the beginning of the random step, size is defined as the population.
2. With respect to each step, size applies the LMS and get its corresponding error value (fitness).

3. A step size having the minimum error select it w.r.t current sample point.
4. With the selected step size LMS applied to get the coefficient value.
5. As the new input sample appears, from the previous generation a new population of step size is created in EP and procedure repeated.

8 Simulation Results

MATLAB 2014b was utilized to implement the modelling and subsequent results have been shown in this section. During modeling, a variety of internal factors of the ACO and ABC algorithm was selected by contemplating the articles [23]. For the channel, the number of plugs chosen for the equalizer is 11 taken over to find out the efficiency of EVSSLMS. The incoming signal includes the 500 examples engendered at random through regular dissemination as demonstrated in Fig. 8. Gaussian noise contains 0.01 standard deviation and zero mean which is combined with the incoming signal as demonstrated in Fig. 2, channel features are provided by the vector:

$$[0.06 - 0.0740.099 - 0.237 - 0.360.11580.360\ 0.2370.0490.099]$$

The ability of the recommended structures was determined in accordance with the conditions of its convergence nature as described in Fig. 9.

This is the randomly generated input signal consists of 500 samples. This signal transfer in a bipolar form (+1, −1). To make the system more complex, random information generated between +1 and −1. This makes the information unpredictable at the receiver side.

Figure 7 illustrates the input signal, and Fig. 8 illustrates engendered incoming signal and signal with noise from the channel.

In Fig. 9, comparative convergence faunas of optimized equalizers EPLMS along with equalizers prompted by ABC and ACO algorithms have been implemented for the various input signal. From the figure, it is clear that the proposed EPLMS performs more iterations with reduced time. Additionally, it was observed that among optimized AEs, EPLMS equalizer is the fastest one.

Figure 10 shows that the EPLMS AEs disclose well BER acts in comparison to the AEs with ABC and ACO specifically the greater SNR area. The comprehensive outcomes have been outlined in Table 1.

From Fig. 11 and Table 2, it has been noticed that the EPLMS equalizer effectively retains its superiority over ACO and ABC triggered AEs in improved communication systems. For illustration, EPLMS equalizers express the BER values of 43.125e−02 for a signal-to-noise ratio value of 8 dB, whereas ACO and ABC generated AEs to offer BER values of 5.058 e−02 and 5.442 e−03 for similar SNR value. More obviously, it was distinguished that the order of improvement in BER behavior of the EPLMS-based EP improved communication system is viewed to be 101 over ABC and ACO based improved communication systems; however, the

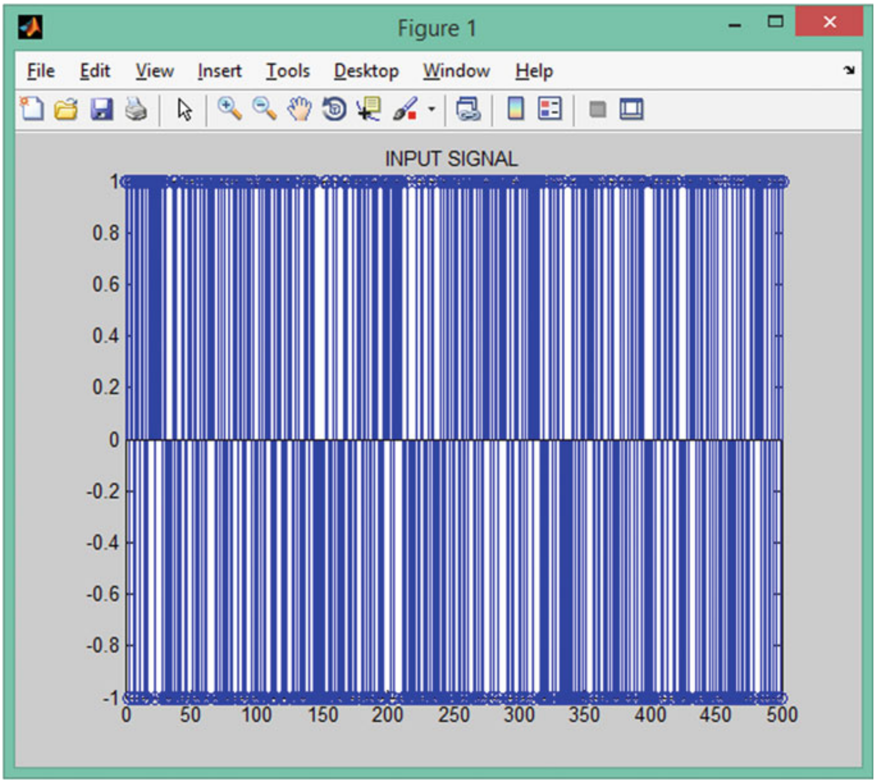


Fig. 7 Input signal for 500 samples

proportion increases in BER value utilizing EPLMS-based improved communication system is experimental to be 23.15%. In connection with this 20 dB SNR level was deemed. Analogous improvements have also been experimental for lower SNR segments. Henceforth, it can remain agreed that EPLMS optimized AEs surpass the AEs synchronized by ACO and ABC methods. However, EPLMS-optimized AE provides the optimal result among all the combinations.

9 Conclusion

Bandwidth-effective data transfer through radio and telephone channels has been made possible through the usage of adaptive equalization to counteract for the dispersal of time launched by the channel. Stimulated by useful applications, a constant research attempt over the past two decades has been producing a wealthy body of fiction in adaptive equalization and the associated more common disciplines of the function of system identification, adaptive filtering, and digital signals. This

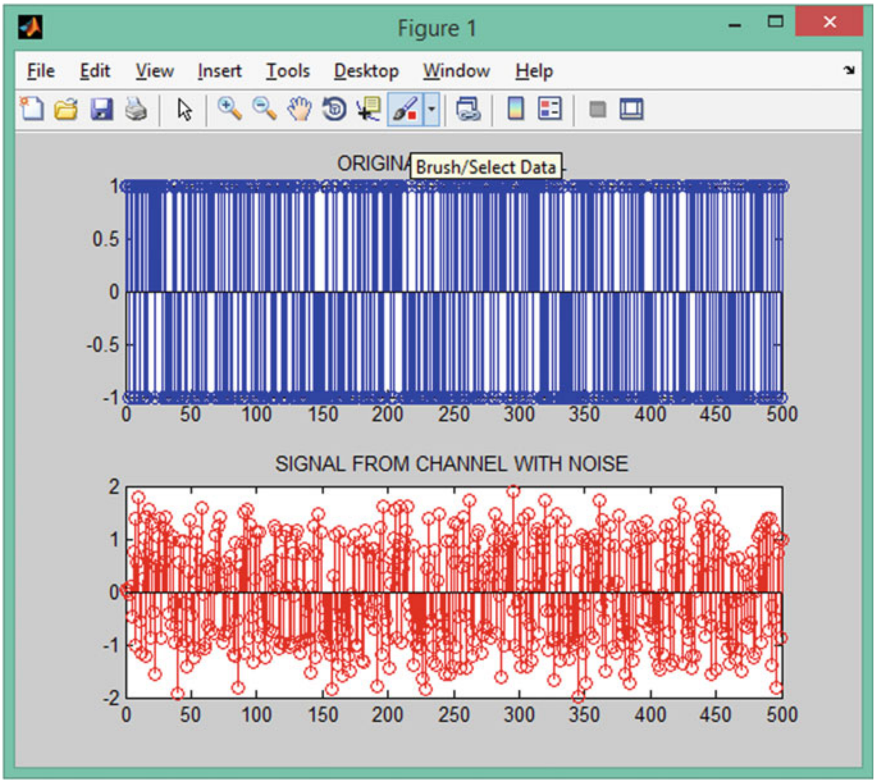


Fig. 8 Engendered incoming signal and signal along with the noise from the channel

article provides a summary of the adaptive equalization. In our design, since evolutionary programming is being used, it will decide what would be the value of step size for a particular application so that mean square error is minimized and convergence is optimal. And also, faster convergence is obtained. Consequently, the effectiveness of an interaction system can be enhanced. A comprehensive review offers that the EPLMS triggered adaptive equalizers to propose quicker merging act in comparison with ABC and ACO structures. Additionally, optimized equalizers are propelled by EPLMS techniques also demonstrate their pre-eminence in accordance with the conditions of BER performance over the ABC and ACO algorithms. Moreover, it has also been observed that EPLMS is better for ABC and ACO optimized AEs. With the layout of the obtaining filters, the impact of Inter-Symbol Interference can be reduced.

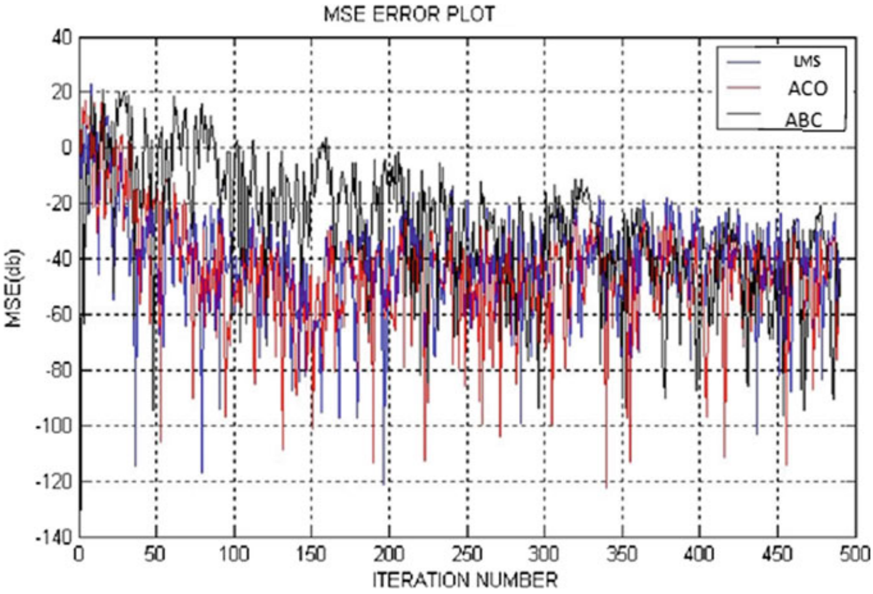


Fig. 9 Comparison of convergence act of LMS with ABC & ACO equalizers

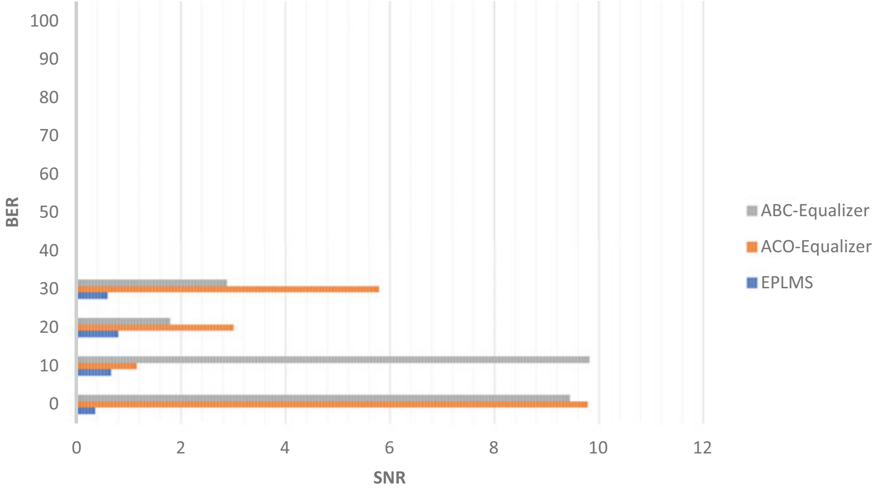


Fig. 10 BER versus SNR analysis of EPLMS optimized equalizer with ABC and ACO equalizers for various input signals

Table 1 BER values input signal

SNR → BER↓	4	8	16	20
EPLMS Equalizer	0.360e−01	0.658e−02	0.799e−04	0.589e−05
ACO-Equalizer	9.784e−02	1.154e−02	3.000e−05	5.789e−06
ABC−Equalizer	9.439e−02	9.820e−03	1.790e−05	2.873e−056

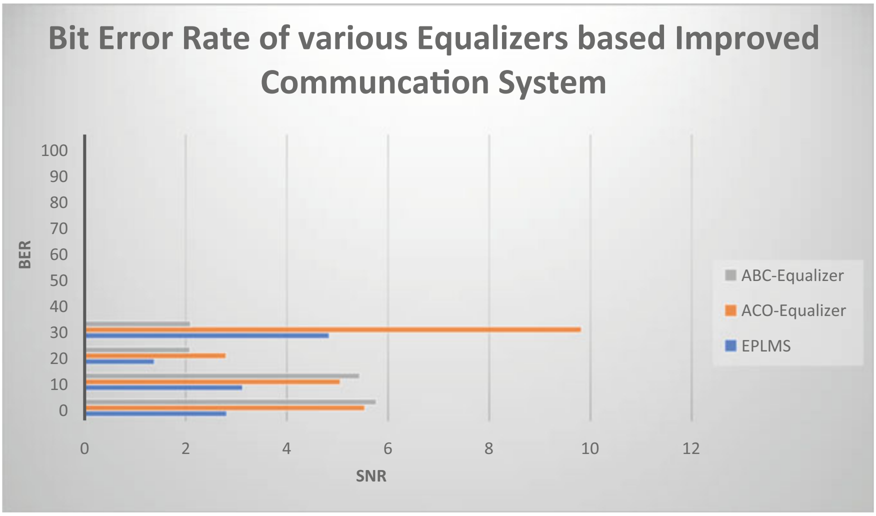


Fig. 11 BER versus SNR analysis acts of the different equalizer-based improved communication system

Table 2 BER values for various optimization algorithms

SNR → BER ↓	4	8	16	20
EPLMS Equalizer	2.811e−03	3.125e−02	1.380e−04	4.837e−08
ACO-Equalizer	5.541e−02	5.058e−02	2.797e−05	9.820e−06
ABC-Equalizer	5.768e−02	5.442e−03	2.085e-05	2.095e−056

References

1. Widrow B, Glover JR, Mc JM. Kaunitz CJ, Williams CS, Hean RH, Zeidler JR, Dong E, Goodlin RC (1975) Adaptive noise cancelling: principles and applications. Proc IEEE 63(12):1692–1716 (December)
2. Harris RW, Chabries DM, Bishop FA (1986) A variable step (VS) adaptive filter algorithm. IEEE Trans Acoust Speech Sign Process ASSP-34(2):309–316 (April)
3. Sethares WA, Mareels IMY, Anderson BDO, Johnson R, Bitmead RR (1988) Excitation conditions for signed regressor least mean squares adaptation. IEEE Trans Circuits Syst

35(6):613–624 (June)
4. Li X, Fan Y, Peng K (2009) A variable step-size LMS adaptive filtering algorithm. In: Proceedings of 5th international conference on wireless communications, networking and mobile computing, IEEE, Sept 2009
5. Wang Y, Bao M (2010) A variable step-size LMS algorithm of harmonic current detection based on fuzzy inference. In: Proceedings of 2nd international conference on computer and automation engineering (ICCAE), IEEE, vol 2, pp 665–668, Apr 2010
6. Xiao Y, Huang B, Wei H (2013) Adaptive Fourier analysis using a variable step-size LMS algorithm. In: Proceedings of 9th international conference on information, communications & signal processing, IEEE, pp 1–5, Dec 2013
7. Schniter P, Johnson CR (1999) Dithered signed-error CMA: robust, computationally efficient blind adaptive equalization. IEEE Trans Sign Process 47(6):1592–1603 (June)
8. Sun L, Bi G, Zhang L (2005) Blind adaptive multiuser detection based on linearly constrained DSE-CMA. IEE Proc Commun 152(5):737–742 (Oct)
9. Al-Awami T, Saif W, Zerguine A, Zidouri A, Cheded L (2007) An adaptive equalizer based on particle swarm optimization techniques. In: Proceedings of 9th international symposium on signal processing and its applications, IEEE, pp 1–4, Feb 2007
10. Boughelala A, Luan X, Leghmizi S (2011) An adaptive channel equalizer using bacterial foraging oriented by particle swarm optimization strategy. In: Proceedings of 3rd international conference on computer research and development, IEEE, vol 2, pp 24–29, Mar 2011
11. Mohammed JR (2012) A study on the suitability of genetic algorithm for adaptive channel equalization. Int J Electr Comput Eng (IJECE) 2(3):285–292 (June)
12. Dey A, Banerjee S, Chattopadhyay S (2016) Design of improved adaptive equalizers using intelligent computational techniques: extension to WiMAX system. In: Proceedings of IEEE Uttar Pradesh section international conference on electrical, computer and electronics engineering (UPCON), pp 305–310, Dec 2016
13. Karaboga D, Basturk B (2008) On the performance of artificial bee colony (ABC) algorithm. Appl Soft Comput 8:687–697
14. Dorigo M, Blum C (2005) Ant colony optimization theory: a survey. Theor Comput Sci 344:243–278
15. Eesa S, Brifcani AMA, Orman Z (2014) A new tool for global optimization problems-cuttlefish algorithm. Int J Math Comput Phys Electr Comput Eng 8(9)
16. Ketonen J, Juntti M, Cavallaro JR (2010) Performance complexity comparison of receivers for a LTE MIMO-OFDM system. IEEE Trans Sign Process 58(6):3360–3372 (June)
17. Paulo SRD (2008) Adaptive filtering: algorithms and practical implementations. Kluwer Academic Publisher, Springer Science and Business Media, LLC. 91
18. Proakis JG (2001) Digital communications. Mc Graw Hill, New York
19. Haykins S (2001) Communication systems. Wiley India Edition
20. Lathi BP (1995) Modern digital and analog communications system
21. Suneel Varma D, Aditya R, Subhashini KR (2013) Synthesis of adaptive antenna with circular geometry employing harmony search and differential evolution techniques. In: 2013 International conference on communication and signal processing, Melmaruvathur, 2013, pp 933– 937
22. Kennedy J (2006) Swarm intelligence. In: Handbook of nature-inspired and innovative computing. Springer, New York, pp 187–219
23. Rao RV, Savani VJ, Vakharia DP (2011) Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems. Comput Aided Des 43(3):303–315

Modified E-Shaped Resonator-Based Microstrip Dual-Mode Bandpass Filter

Shobha I. Hugar, Vaishali Mungurwadi, and J. S. Baligar

Abstract An modified E-shaped dual-mode resonator is proposed to design two types of dual-mode bandpass filters such as **Filter I** and **Filter II**. The proposed resonator comprises of a half-wavelength hairpin resonator and a meander ring on its symmetrical plane. As the structure is symmetrical, it supports odd and even mode theory. In **Filter I**, based on coupling gaps between feedlines and middle arm of E-shaped resonator, the location of upper stopband transmission zero is controlled. Using two tuning stubs incorporated with inner meander ring, passband center frequency and bandwidth are controlled in **Filter II**. Additional transmission zero is procured in upper stopband using source load coupling scheme to improve the selectivity of the filter. Both filters are designed for passband center frequency 2.5 GHz and fractional bandwidth 0.4.

Keywords Dual-mode · Fractional bandwidth · Transmission zeros · Quality factor · Radio frequency (RF)

1 Introduction

Dual-mode microstrip bandpass filters have been used in front end wireless communication system due to their significant features such as small size, good passband performance, high-quality factor, and low loss. The major limitations of single-mode filters such as narrow bandwidth and low-performance characteristics have made the researchers develop dual-mode filters for radio frequency (RF) and microwave applications.

S. I. Hugar (✉) · J. S. Baligar
Dr. Ambedkar Institute of Technology, Bangalore, Karnataka, India
e-mail: Shobha_hugar@yahoo.co.in

J. S. Baligar
e-mail: jbaligar@gmail.com

V. Mungurwadi
Visvesvaraya Technological University, Bangalore, India
e-mail: Vaishalibm18@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_49

E-shaped resonator was initially proposed in [1], where it was designed by two quarter-wavelength resonators with a quarter-wavelength open stub at center acting as K inverter to control coupling strength and to create transmission zeros at the desired frequency. Doublet and extended doublet coupling schemes were proposed in [2] to design two-pole and three poles bandpass filters, respectively, using E-shaped resonator. But in these filters, the selectivity was improved at the cost of larger filters size. To realize additional transmission zero in upper stopband [3], a capacitor C introduced with source load coupling and location of transmission zero was controlled by adjusting the value of C.

Stepped impedance E-shaped hairpin resonator loaded with T-shaped open stub was proposed to design dual-mode dual-band BPF [4] with wide stopband. Dual-mode ring [5] embedded with quarter-wavelength split open-end resonator was proposed to form extended doublet coupling which has improved the selectivity of the filter by creating one pair of transmission zeros near passband. Center frequency and bandwidth tuning were proposed in [6] by changing the finger width of inter-digital loading element. It is observed that filters reported in [3, 5] have a narrow upper rejection band.

An modified E-shaped dual-mode resonator is presented to design two types of dual-mode bandpass filters **Filter I** and **Filter II**. The proposed resonator comprises of a half-wavelength hairpin resonator and a meander ring on its symmetrical plane. As the structure is symmetrical, it supports odd and even mode theory. The gap between feedlines and mid-arm of E-shaped resonator controls location of transmission zero in upper stopband in **Filter I**. Using two tuning stubs, the passband frequency and bandwidth are controlled in **Filter II**.

2 Modified E-Shaped Dual-Mode Resonator

Figure 1a shows a layout of proposed modified E-shaped dual-mode resonator BPF with source load coupling scheme on a substrate with relative permittivity 10.2 and thickness 0.635 mm. The resonator comprises $\boldsymbol{\lambda_g/2}$ a hairpin resonator with 50 Ω impedance and a meander ring on its symmetrical plane. $\boldsymbol{\lambda_g}$ is the guided wavelength at center frequency 2.5 GHz. The dimensions of the proposed resonator are L_1 = 7.18 mm, L_2 = 12.77 mm, L_3 = 1 mm, L_4 = 1.8 mm, W = 0.59 mm, W_1 = 0.4 mm, W_2 = 0.2 mm. Figure 1b represents a source load coupling scheme where the resonator 1 represents odd mode frequency and resonator 2 represent even mode frequency. The input source is coupled to both modes represented by dark lines and so is the output. The dashed line represents the weak coupling between input and output. The coupling matrix for a given source load coupling scheme can be written as [3] (Fig. 2).

The proposed modified E-shaped dual-mode resonator has the following interesting properties.

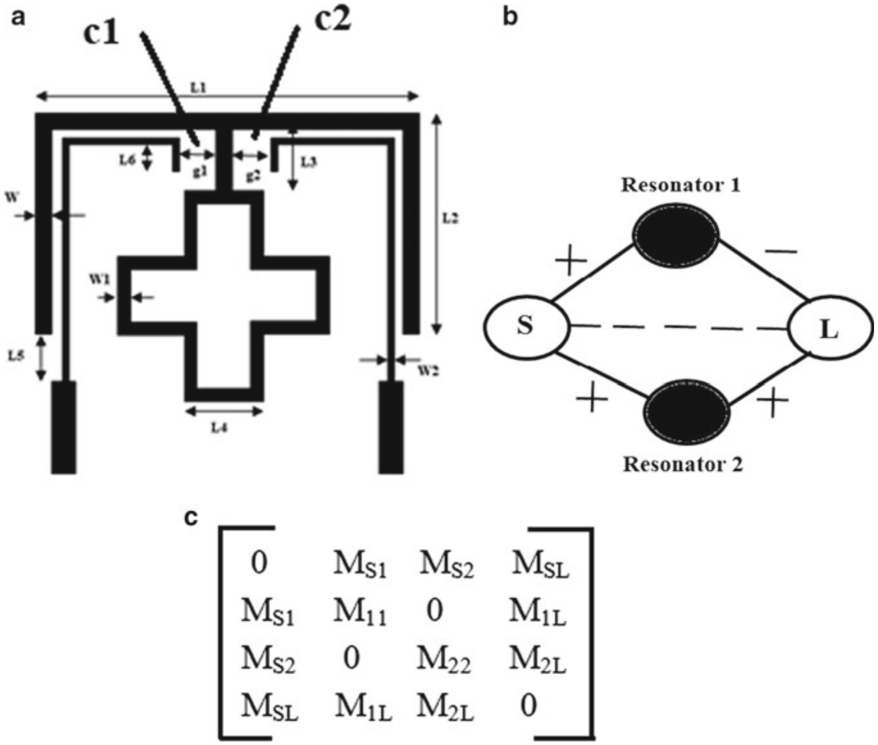


Fig. 1 **a** The layout of the proposed filter **(Filter I)**. **b** Source load coupling scheme. **c** Coupling matrix [3]

(a) The coupling strength between external feeding circuit and odd mode is greater than even mode and hence $\mathbf{M_{s1} > M_{s2}}$
(b) The structure has finite inherent transmission zeros since $\mathbf{M_{s1}}$ and $\mathbf{M_{s2}}$ are not equal.
(c) As $\mathbf{M_{11}} > 0$ and $\mathbf{M_{22}} < 0$, the transmission zeros are greater than zero and appear on the real axis.
(d) Since $f_o > f_{\mathbf{odd}}$ and $f_o < f_{\mathbf{even}}$, the transmission zero appear in the upper stopband.
(e) Two degenerative modes will not couple.

2.1 *Filter I*

Filter I designed using modified E-shaped dual-mode resonator shown in Fig. 1a exhibits genetic transmission zero in upper rejection band. The proposed filter is designed for passband center frequency $f_o = 2.5$ GHz and fractional bandwidth (FBW) of 0.4. Here, two capacitors C1 and C2 are introduced with source load coupling to control the location of additional transmission zero TZ_2 in the upper

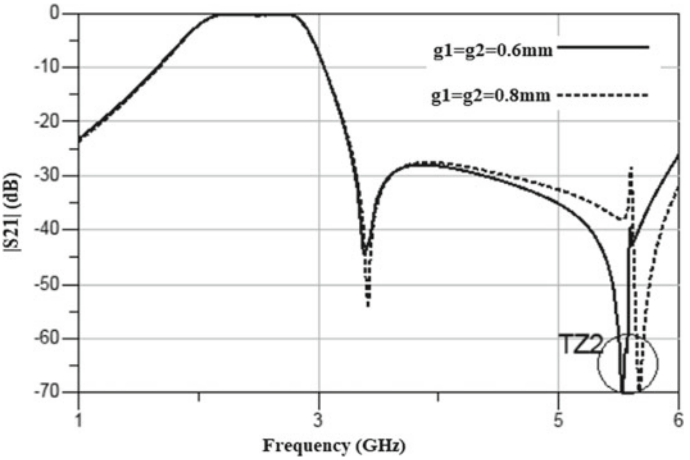


Fig. 2 Change in location of transmission zero TZ_2 with increasing coupling gaps g_1 and g_2

stopband. It is noted that with an equal increase in coupling gaps g_1 and g_2, the equivalent capacitance is reduced which shifts transmission zero TZ_2 to higher frequency location resulting into wider stopband. Figure 2 shows a change in the location of transmission zero TZ_2 with increasing coupling gaps g_1 and g_2. From EM simulation results, it is noted that the designed filter has S21 < 1 dB and S11 > 10 dB in the passband and wide rejection band up to 5.9 GHz with 30 dB attenuation level.

2.2 Filter II

The layout of **Filter II** designed using modified E-shaped dual-mode resonator is shown in Fig. 3. The dimensions of the filter are the same as in Fig. 1a. As the structure is symmetrical, it supports odd and even mode theory. Here, additional two tuning stubs of length L are placed on the inner meander ring to tune even mode resonance frequency. Increase in the length L of two tuning stubs increases the susceptance [7] which shifts even mode resonant frequency and upper passband edge to lower frequencies. From simulation results, it is also noted that the shift of upper passband edge to lower frequency results in narrow passband which increases selectivity and Q of the filter. Figure 4 depicts the shift in even mode resonance and upper passband edge to a lower frequency for various lengths L of tuning stubs. It is observed that these shifts in even mode resonance frequency and upper passband edge results in the tuning of both passband center frequency f_o and bandwidth.

The proposed filter has $S_{21} < 1$ dB and $S_{11} > 10$ dB in the passband for all values of L. From the simulation results, it is noted that with $L = 0$ (without stubs), 1 and 1.5 (in mm) the even mode frequency is at 2.75 GHz, 2.64 GHz, 2.53 GHz, and upper

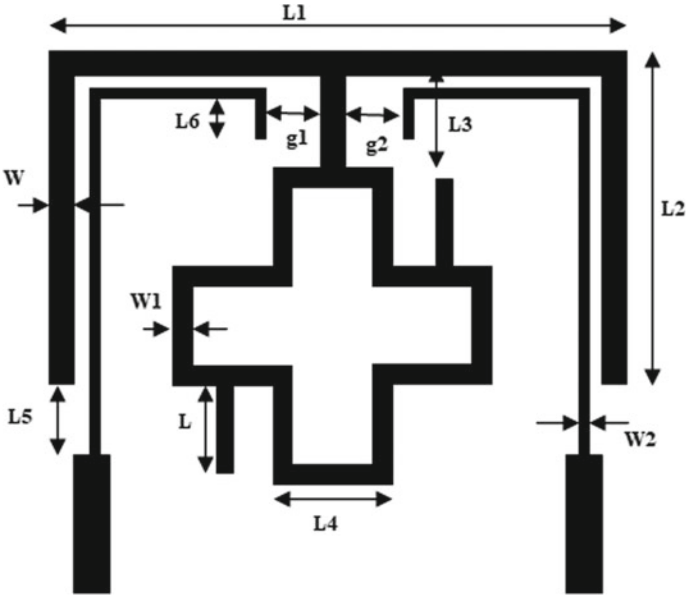


Fig. 3 Proposed dual-mode BPF with two tuning stubs (Filter II)

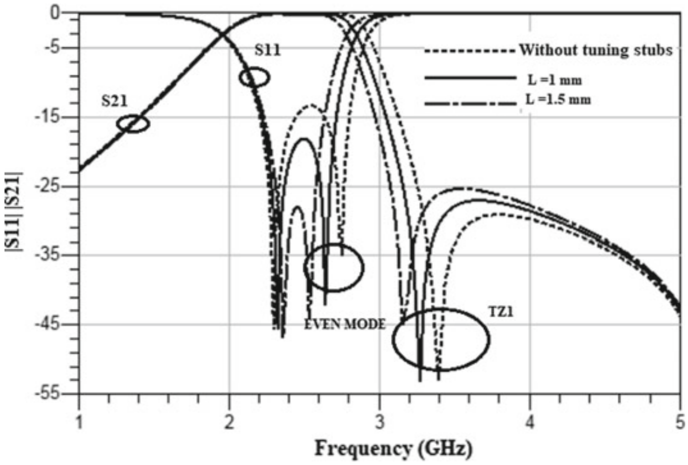


Fig. 4 The shift in even mode resonance and upper passband edge to a lower frequency for various lengths *L* of tuning stubs

passband edge is at 3.39 GHz, 3.27 GHz, 3.15 GHz, respectively. With $L = 1.5$ mm, the center frequency is tuned from 2.5 to 2.45 GHz and FBW from 0.4 to 0.33.

3 Conclusion

Two types of dual-mode bandpass filters **Filter I** and **Filter II** are designed using modified dual-mode E-shaped resonator at a center frequency 2.5 GHz and fractional bandwidth 0.4. To achieve desired response, first characteristics of the proposed resonator are discussed. In **Filter I**, the location of additional transmission zero TZ_2 is controlled by varying the gaps between feedlines and mid-arm of E shaped resonator. It is noted that wider gaps shifts TZ_2 to higher frequency resulting in a wider rejection band. Both passband center frequency and bandwidth are controlled in **Filter II** using two tuning stubs. The special advantage of **Filter II** is that with an increase in the length of tuning stubs resulted into narrow passband which is important for achieving high selectivity and good Quality factor(Q).

References

1. Lee J-R, Cho J-H, Yun S-W (2000) New compact bandpass filter using microstrip/4 resonators with open stub inverter. IEEE Microwave Guided Wave Lett 10(12) (Dec)
2. Liao C-K, Chi P-L, Chang C-Y (2007) Microstrip realization of generalized Chebyshev filters with box-like coupling schemes. IEEE Trans Microwave Theory Tech 55(1) (Jan)
3. Zhang X-C, Yu Z-Y, Xu J (2008) Design of microstrip dual-mode filters based on source-load coupling. IEEE Microwave Wirel Compon Lett 18(10) (Oct)
4. Wang J, Ge L, Wang K (2011) Compact microstrip dual-mode dual-band bandpass filter with wide stopband. IEEE Electron Lett 47(4)
5. Zhou M, Tang X, Xiao F (2010) Miniature microstrip bandpass filter using resonator-embedded dual-mode resonator based on source-load coupling. IEEE Microwave Wirel Compon Lett 20(3) (Mar)
6. Karpuz C, Özdemir PO, Fırat GB (2016) Design of fourth order dual-mode microstrip filter by using interdigital capacitive loading element with high selectivity. Proc Eur Microwave Conf
7. Kundu AC, Awai I (2001) Control of attenuation pole frequency of a dual-mode microstrip ring resonator bandpass filter. IEEE Trans Microw Theory Tech 49(6) (June)

An Analysis of Air Quality by Vehicular Exhausts

V. Kanpurrani and A. L. Vallikannu

Abstract The level of emission of air pollutant contents is to be monitored to improve the life of the vehicles and improve the quality of the air in the environment. The automatic intelligence system is a dire need to monitor the emission of air pollutants of vehicles, particularly in the smart cities. The proposed system is to monitor the air pollutants in the air and generates the alert for the remedial action when the air pollution exceeds the normal level. The proposed idea has experimented in the simulated smart city environment. The smart city environment is simulated with wireless communications with smart sensors for measuring the air quality, the cloud storage is used for data storage and processing. The experimental results have been recorded and analyzed. The accuracy has been computed and discussed. The proposed system has yielded as high as 11.67% more accuracy than the existing system.

Keywords Radio-frequency identification (RFID) · Internet of things (IoT) · Liquid crystal display (LCD)

1 Introduction

Air pollution control tools are referring to regulate and eliminate the emission of potentially hazardous substances of particulate of gases and matter. The pollution control system is widely understood as air-pollution control system. There are a variety of pollution control systems that focus on different sectors like the air pollution control system and water pollution control system. The pollution control system helps reducing or to prevent polluting particles of hazardous particles to get directly into the environment system.

Nowadays, air pollution control system is a very important concern for industries as many industries releasing toxic waste in the air for that reason every industry

---

V. Kanpurrani (✉) · A. L. Vallikannu
School of Electrical Sciences, Hindustan Institute of Technology and Science, Chennai, India
e-mail: kanpurrani@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_50

needed to release any kind of air after purification which helps in reducing pollution in the air.

1.1 Major Types of Air Pollution Control

(a) *Particulate Control*: Special machinery is used to filter out particulate pollution from gases. It is a kind of physical matter separation from air.
(b) *Ways of particulate control*: Electrostatic precipitators, cyclone separators, fabric filters
(c) *Gas Control*: In normal chemical methods are used for the separation of pollutant components from gas. It is always useful to remove hazardous gases from waste or else it may lead to acid rain.
(d) *Ways of particulate control*: Scrubbers method, incineration method, carbon capture.

The modeling of air pollution is dependent on the transport and turbulent mixing of emissions. One of the major contributors to pollution is vehicle emissions. Emissions that are released directly from the cars and trucks are the primary source of vehicular pollution. Motor vehicles also pollute the air during the processes of manufacturing, refueling, and from the emissions associated with oil refining [1]. BS are emission standards by the Government of India to regulate the output of air pollutants from combustion engine equipment, motor vehicle. To control the pollution exhausted by vehicles, the amount of air pollution is needed to be calculated, and vehicles causing pollution must be identified. Internet of things may become helpful in cities for monitoring air pollution from vehicles, and the amount of pollution can be gathered and analyzed. This system is specially designed to operate the system using a sensor network and gather information about pollutant levels discharged by the vehicles. To control the emission from the vehicle, the Government has implemented a PUC certification. PUC certification is compulsory for all vehicles on Indian roads. It is to ensure that the vehicle is under the norms of pollution. The validity of the certificate is 6 months. After that, a new certificate must be taken. To overcome this problem, our proposed system has an inbuilt system inside the vehicle that checks the continuous emission values.

2 Literature Survey

The major source of air pollution in the twenty-first century is the road vehicles. The transportation of vehicles exhausted 50% of the carbon monoxide in the air. The increase in population and the busy world, the graph of usage of the personal vehicles in the cities increases in an exponential manner over a decade.

The air pollutant particles include carbon monoxide, nitrogen oxides (NOx), and hydrocarbon (HC) are rapidly increasing in levels in the air. The increase in air pollutants decreases air quality [2]. It also affects the health of the public. The lowest air quality makes people hard to breathe and yields many diseases include cancers. The increase in urbanization and industrialization produces the bad quality of the air. It is very difficult to monitor the vehicles to control air pollution by vehicles. Most of the vehicle users are ignoring to maintain the quality of the vehicles and hence the air quality gets affected.

Though the government has formed many standards and mechanisms to monitor and control air pollution through transportations, controlling air pollution is still not achieved. The ignorance of the vehicle user and the size of populations are the greatest challenges in monitoring vehicle emissions [3]. The dire need of today's air pollution controlling in transport vehicles is an intelligence system that can monitor the air emission particle of the vehicles and analyze the measured data and inform the concerned authorities to take the necessary steps to protect the movement of the vehicles. There are different approaches developed for monitoring vehicular pollution. For instance, in Rushikesh et al. [4] the concept of talking about is to control the pollution level with the help of MQ sensors, RFID, and Arduino. If the pollution level is above the threshold, the authorities will be informed about it. The microcontroller then reads the level and it will be sent to the vehicle owner and the server for future analysis. In Manna et al. [5], the high traffic area is selected to monitor the pollution level with the help of RFID placed at a short fixed distance along with the sensor nodes at the roadside. The tag number of the vehicle is identified by an radio-frequency identification (RFID) reader and is transmitted to the server. When the sensor level is high than the threshold, a message is sent to the owner. According to Priyanka et al. [6], the sensors are placed at the outlets of the vehicles, and the output of it is given to the input of Arduino, and if beyond the threshold value, the information is sent through GSM to IBM Watson internet of things (IoT) platform [7]. The data collected from each of the vehicles are viewed by logging in using user credentials on the dashboard. Considering Pavani et al. [8], the gas sensor used for checking the pollution level is the TGS4161 electrolyte carbon dioxide sensor which needs to be calibrated by exposing them to several concentrations of gas and each sensor output is then plotted accordingly. The goal of Rahman et al. [9] is to interface the output from the sensors used with the online development environment Thing Speak. The signal inputs include temperature variations, sound, potentiometer readings, and so on. It represents the output data as a graphical format as well as in an excel sheet.

3 Proposed System

The proposed work has been designed by sensor module, communication module, cloud storage, learning algorithms, cloud database, and learning algorithms as shown in Fig. 1. The proposed system has experimented in the simulated defined environ-

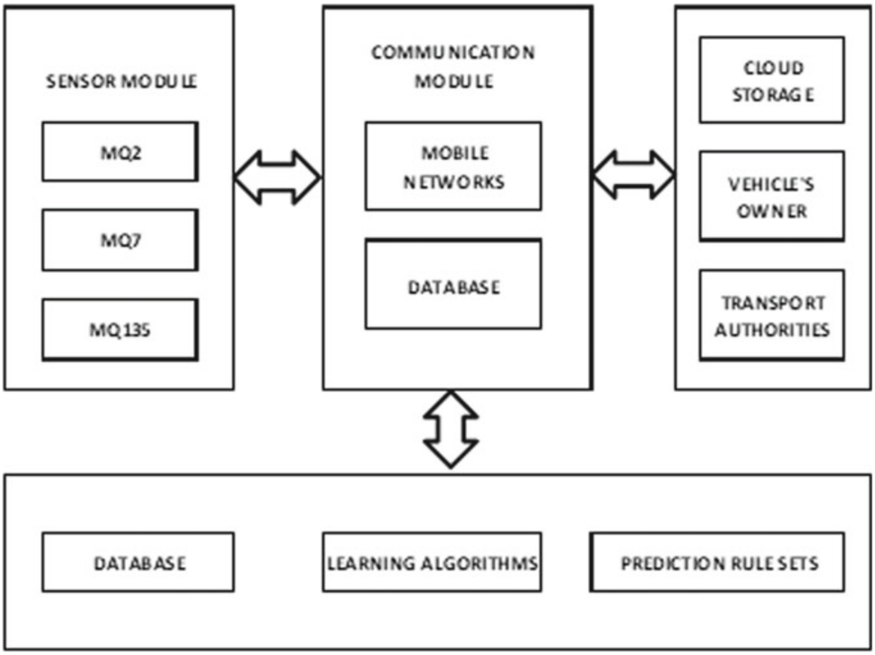


Fig. 1 The block diagram of the air pollution monitoring system

ment. The wireless module was used for communication as in smart city environment [10, 11].

3.1 Vehicle

The vehicle consists of sensors, a controller, a buzzer, and an liquid crystal display (LCD). The sensor has collected the data of pollution, and the level of pollutants has been forwarded to the controller [12]. If the level has exceeded the threshold value, the alert has generated and displayed in the LCD. The controller contains the wireless connection used to send the information to the owner of the vehicle to make him aware of the condition of the vehicle. By collecting the information about the conditions of the vehicular, data has been analyzed from the user-level application. Internet SMS is sent to the vehicle owner by analyzing the data from the application for the use [13].

3.2 Sensors

It measures the emitted values of the exhaust gases and the level of respective gases, like hydrocarbon (mq2), carbon monoxide(mq7), and nitrogen oxide (mq135). The level of gas values has been processed to the algorithm for storage and communication to the cloud storage [14].

3.3 Communication Module with Cloud

It consists of a database and wireless communication, the processing of communication, the sensed data has stored in the database. The stored data has a reference number to be processed with the threshold values of the spatial locations. The reference data has been used for information retrieval between the cloud and the database. The data consist of the level of pollutants, temperature, and id of the owner's vehicle [15]. These databases have been stored in the cloud storage. The cloud is acting as a centralized database for storage and processing. The server will get the information from the vehicle when it attains the threshold value of the level of pollution.

3.4 Processing Module

The processed module consists of a database, learning algorithms, and prediction rule sets. The prediction rule sets are used to predict the near future forecasting to make an advance alert to the users. The sensed data value of the pollutants is compared with threshold value as defined. If the sensed data value exceeds the threshold value, the alert is generated. The generated alert is communicated to the users and other competent authorities for remedial action. The learning algorithm is illustrated as follows.

```
Pseudocode: Vehicle Pollution Monitoring System – Generating Alert
Input Environment = Transport Vehicle in Smart City
Sensors Used = mq2, mq7, mq135
Time Interval to Measure = 30 min
Communication = Wireless /Mobile Network
Storage space = Data Server, Cloud Setup

For each time interval 't'

1. The mq2, mq7, mq135 sensors measures the level of gas carbon monoxide (CO),
   nitrogen oxides (NOx), hydrocarbon (HC).
2. The measured data is recorded and compared with the threshold values.
```

3. *If the trend of measured data is likely to fall into the threshold value of the geographical locations, the alert is generated and sends to the owner of the vehicles.*
4. *If the measured data is exceeded the threshold value, the alert is generated and send to the owner of the vehicles and the concern authorities.*
5. *Else if Step 2 to Step 4 is repeated in a looping fashion.*

The proposed system has experimented in the simulated smart city. The smart city is simulated with wireless communications and a common cloud for data storage. The government authorities include transport, traffic control, and mobile communications are also simulated in the defined smart city.

4 Result and Analysis

The accuracy of alert is considered as performance parameters, it is defined as

$$\text{Accuracy of Alert} = \frac{\text{No.of\_ Actual\_ Alerts\_ of\_ 100\_ trails}}{\text{No.of\_ Expected\_ Alerts\_ of\_ 100\_ trails}}$$

The experiments are carried out independently for 100 trails and the data is recorded. The average is taken after ignoring the outline's values. The analysis was carried out under four cases as True-True, True-False, False-True, False-False cases as defined in Table 1.

The presented work analyzed the pollution control data in four scenarios, the alert conditions were tabulated in Table 1. The interactive results of predicted accuracy have yielded better than the actual values of various scenarios conditions of the variables.

The accuracy of alert conditional values has tabulated in Table 2. The proposed work yielded a better performance as 98.1% for False–False case as shown in Table 2. However, the proposed work could not get better performance for True–False case as shown in Fig. 2. The accuracy of actual values for the four categories as True–True, True–False, False–True, False–False as shown in Fig. 2.

The results are recorded for prediction analysis. The accuracy on prediction is plotted as shown in Fig. 3. The proposed work was yielded the better performance as 97.98% for False–False case.

Table 1 Recorded scenarios

S.NO	Condition	Alert
1	T_c and T_c	The threshold value has exceeded
2	T_c and F_c	Not generated and the actual data exceed threshold values
3	F_c and T_c	Generated and the actual data exceed threshold values
4	F_c and F_c	Not Generated and the actual data is not exceeded threshold values

Table 2 Accuracy of alert generation on exceed the threshold values

S.NO	Condition	Accuracy (%)
1	T_c and T_c	61.76
2	T_c and F_c	43.21
3	F_c and T_c	71.78
4	F_c and F_c	98.10

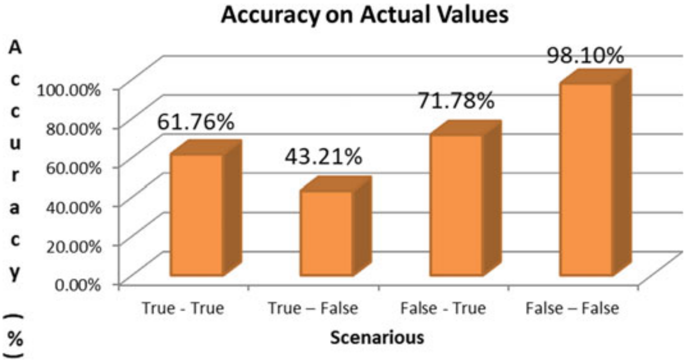


Fig. 2 Accuracy on actual values for alert generation

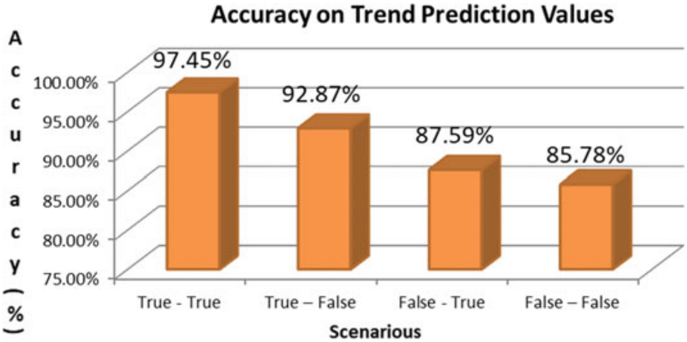


Fig. 3 Accuracy of alert generation on trend prediction

The proposed system is performing well as 97.45% in the True-True scenario. The performance was 92.87%, 87.59% and 85.78% for True-False, False-True, and False–False, respectively. It proved that the proposed work is suitable for the accuracy on prediction while monitoring the pollutants in the air quality.

5 Conclusion

It is a dire need to monitor air pollution in the city to maintain the quality of air. The proposed system has been designed to monitor the air quality in the city. The sensor module has measured the air quality level. The measured values are compared with the threshold values and generate the alert as needed. The learning algorithm plays a vital role in predicting the pollution level shortly. The proposed system has yielded a better performance as 99.61%, 98.10% for prediction, real-time, respectively. The future direction of the work shall focus on the real-time measurement of pollutants in the air.

References

1. Chiti F, Fantaccia R, Gu Y, Han Z (2017) Content sharing in internet of vehicles: two matching-based user-association approaches. Veh Commun 8:35–44 (Apr)
2. Contreras-Castillo J, Zeadally S, Guerrero-Ibáñez JA (2018) Internet of vehicles: architecture, protocols, and security. IEEE Internet Things J 5(5):3701–3709 (Oct)
3. Duangsuwan S, Takarn A, Nujankaew R, Jamjareegulgarn P (2018) A study of air pollution smart sensors LPWAN via NB-IoT for Thailand Smart Cities 4.0. In: 2018 10th international conference on knowledge and smart technology (KST), Chiang Mai, 2018, pp 206–209. https://doi.org/10.1109/kst.2018.8426195
4. Korunoski M, Stojkoska BR, Trivodaliev K (2019) Internet of things solution for intelligent air pollution prediction and visualization. In: IEEE EUROCON 2019—18th international conference on smart technologies, Novi Sad, Serbia, 2019, pp 1–6. https://doi.org/10.1109/eurocon.2019.8861609
5. Kularatna N, Sudantha B (2008) An environmental air pollution monitoring system based on the IEEE 1451 standard for low cost requirements. Sens J IEEE 8:415–422. https://doi.org/10.1109/JSEN.2008.917477
6. Divya A, Kiruthika R, Gayathri D (2019) Detecting and analysing the quality of air using low cost sensors to reduce air pollution in urban areas. In: 2019 IEEE international conference on system, computation, automation and networking (ICSCAN), Pondicherry, India, 2019, pp 1–5. https://doi.org/10.1109/icscan.2019.8878780
7. Kavitha BC, Vallikannu R (2019) IoT based intelligent industry monitoring system. In: 2019 6th international conference on signal processing and integrated networks, SPIN 2019, 8711597, pp 63–65
8. Malky S, Kostanic I, Altheiab K, Alharbai W (2019) Evaluation of precalibrated electrochemical gas sensors for air quality monitoring systems. In: 2019 international conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData), Atlanta, GA, USA, 2019, pp 967–973
9. Postolache OA (2009) Smart sensors network for air quality monitoring applications. IEEE Trans Instrum Meas 58(9) (September)
10. Xu Y, Chen X, Liu A, Hu C (2017) A latency and coverage optimized data collection scheme for smart cities based on vehicular ad-hoc networks. Sensors 17(4):888
11. Hemanth J, Fernando X, Lafata P, Baig Z (2018) International conference on intelligent data communication technologies and internet of things
12. He Y, Sun D, Zhao M, Cheng S (2018) Cooperative driving and lane changing modeling for connected vehicles in the vicinity of traffic signals: a cyber-physical perspective. IEEE Access 6:13891–13897

13. Silva CM, Silva FA, Sarubbi JFM, Oliveira TR, Meira W Jr, Nogueira JMS (2017) ‘Designing mobile content delivery networks for the Internet of vehicles’. Veh Commun 8:45–55 (Apr)
14. Meneguette RI, Boukerche A (2017) SERVitEs: an efficient search and allocation resource protocol based on V2V communication for vehicular cloud. Comput Netw 123:104–118 (Aug)
15. Rani VK, Sudalai MT (2015) An innovative design of intelligent cradle for infants. Int J Appl Eng Res 10(8):20867–20873

Secured Data Encryption and Decryption for Transmission in Device Independent Systems

B. N. Lakshmi Narayan, N. H. Prasad, and A. Mariyan Richard

Abstract In general, a company will have its branches at different locations and every branch will have different hardware and software configurations. Different configurations have different issues in transferring the data from various branches to corporate office and vice versa. To overcome the challenge, this application remains as an alternative solution. Data transferred through the network remains insecure due to its plain text format. Regarding the data restoration, it becomes incompatible for software as well as hardware. For security concern, the security process handling in data encryption and data decryption using cryptic service provider for RSA encryption and decryption using public and private key. The keys were stored in server as a XML format. This encryption is an asymmetric encryption. This data also accessed through device independent system like mobile, PDA, etc. The quality of file to trace in progress prominence of the file which users works on and after getting the files, the file will be sent to quality analysis person and it will be obtained back with the status before restoring a database. The quality of file will be checked and sent back to the administrator. The unsatisfactory files are then sent back to client for rectification.

Keywords Data Encryption · Triple Data Encryption · RSA · Asymmetric Encryption

1 Introduction

The paper entitled in secured data transmission in device independent system which deal regarding the issue of data security. The conventional method of data is plain text format. This method will raise the security issue of data when transfer from

---

B. N. Lakshmi Narayan (✉) · N. H. Prasad · A. Mariyan Richard
Department of MCA.NMIT, Bengaluru, India
e-mail: narayan614@gmail.com

N. H. Prasad
e-mail: naikphd@gmail.com

A. Mariyan Richard
e-mail: mariyanrich01@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_51

source to destination. Any social networking site user, example a Facebook user can login to his account by giving his user id and password. If the user is authorized, only then he would be allowed to get his Facebook home page.

The administrator to get files from the client and register the file information, depending on the information, the encrypted file was decrypted and stored into server database. The restored data was fetched from database and processed by administrator. If the file to be unsatisfactory, then the file is send back to the client for correction. The security process was handling data encryption and data decryption using cryptic service provider.

This paper consist of following module.

- Authentication Module
- Database Selection
- Key Generation Module
- Encryption
- Decryption

Generally, the authentication module deals with the process of authentication. A Facebook user will get his Facebook homepage only after he has cleared the process of authentication [1]. Unauthorized persons were automatically redirected into authentication area [2]. This module provides to create new user registration, change their password, and recover password when they are forget it. The server deals with the admin side [3]. The administrator gets files from the clients and registered to it. The files are in encrypted XML format [4]. The files are decrypted and data will be restored into server database [5]. The administrator fetches data from database and converted into machine independent language then uploaded into server. The client module deals with the client side [6]. The client fetches data from database and provides security to that data and converted into device independent language. This data was uploading to the server for administrator verification [7]. The security deals with handling data encryption and data decryption using cryptic service provider. The public and private keys will be created and stored in server as a XML format [8]. The public key and private key was created and stored into Internet information server [9]. The keys used when security was applied. The conversion module deals with the data convert XML format into plain text format vice versa [10–13]. This conversion was taken place in server and client side. The quality analysis module deals with the data verification of each file before restores [14]. The files are sending to quality analysis persons for data verification. The analyzer was analyzed to it and sends back the file with status.

This module helps us to track the data through various devices such as mobile phone, PDA, etc.

(A) *Demerits Of The Existing System*

- The data transfer is a plain text format.
- Due to plain text, the data will be insecure
- Bulk data will not transfer from source to destination

- The database structure will be copied and restored
- The system occupy huge memory space
- Possible of missing data
- The version of software is incompatible

2 Proposed System

To provide security when the data was transfer in an insecure channel and convert machine or device independent language for accessing the external devices or systems.

The proposed system is a computer-based system which overcomes the existing system. This proposed system provided to fetch all the data from database and converted into encrypted XML format. This XML format do not dependent any operating system, mail server, browser, server, and any other hardware. So, this data was software and device independent to the systems.

(A) *Benefits of Proposed System*

- While fetching data, the database structure will not copied.
- The Rivest–Shamir–Adleman (RSA) Encryption / Decryption standard was used for security.
- The data will be converted into XML format.
- The XML file will be checked for quality analysis person before restore database.
- Movement of data will be traced by both server side and client side persons.
- The data will be accessed through device independent system.

3 System Design

This paper has the process of encryption and decryption methods to send the data from one place to another place. To proceed this process, we have to decide which data has to be sent. After the selection of database, we have to choose the algorithm to generate keys to encrypt and decrypt. The following sections define the algorithms which are used in this paper.

(A) *RSA*

RSA is a one of the first public-key cryptosystems algorithm is used to generate keys for below signature.

- Public Key: (n,e)
- Private Key: (n,d)

(B) *RIJNDAEL/AES*

Rijndael is a symmetric algorithm. The rijndael is also called advanced encryption standards (AES). This has the key size of 128,192, or 256 bits (32 bytes). The keys here are.

- Password,
- Secret Key.

(C) *Ron Rivest Cipher2 (RC2)*

Ron Rivest who designed RC2 replaced DES with his designed cipher. The designed cipher recorded the speed which was 3 times more compared to DES. The block size of input and output are 64 bits each. The key size of the block differ from one byte and it may reach up to 128 bytes. The proposed system and its implementation are used up to eight bytes. The proposed algorithm is designed to make a firmware for 16-bit microprocessors. The keys are.

- Password,
- IV (Initialization Vector)

(D) *Data Encryption Standards (DES)*

The encryption and decryption of electronic data goes to as the input of 64 bit blocks into DES with 64 bits as cipher text. The key length of DES is 56 bit. But, it can accept the input key of 64-bit long. *DES works on bits (0′s and 1′s), where group of four bits forms hexadecimal number.*

Example: (i) 1111 = F.

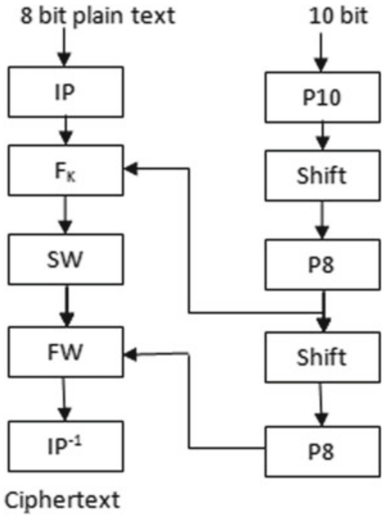
(ii) 0011 = 3.

DES works on bits or binary numbers by encrypting to digital computers. Hexadecimal format is used to encrypt the 64 message bits. DES takes 64 bits of data into each set by converting it into two blocks of 32 bits each.

(a) L = Left half block
(b) R = Right half block.

DES has majorly features into five functions.

(a) Initial Permutations (IP)
(b) Complex functions (FK), which involves both real and imaginary arguments and it produces same structure of complex types.
(c) A combination of one of its permutation function which switches (SW) the controls from one block to other block (i.e., L and R).
(d) Switch Function (SF).
(e) Inverse Initial Permutation (IP-1).



where

Ciphertext = IP-1(Fk2(SW(Fk1(IP(Plaintext))))).

And Key K1 = P8 (Shift(P10(key))).

K2 = P8(Shift(Shift(P10(key)))).
Example: Let P be the plain text message of variable M = 10,111,010, K represents as a key with value of K = 1,001,101,011.
DES is itself is fairly divides the key into 10-bits between the sender and receiver.
After receiving the main key, it generates two other sub keys of size 8 bits each.
First, let the 10-bit key is initialized as follows.
K10 (k1, k2, k3, k4, k5, k6, k7, k8, k9, k10).
And permutation P10 is initialized as below:
P10 (k1, k2, k3, k4, k5, k6, k7, k8, k9, k10) = (k3, k5, k2, k7, k4, k10, k1, k9, k8, k6).
So, we consider the initial permutation combination using the IP function as below:
IP (1 2 3 4 5 6 7 8) = (2 6 3 1 4 8 5 7).
So, according to IP, the plaintext gives the output as below.
IP: 11,010,101.
Inverse the initial permutation, as result.
P-1 = 10,101,011.

K = 000,101,111.

CS1 = 00,001 11,111.

CS2 = 10,000 11,111.

As a result of the permutation, we define the mapping variable F, where it takes the 4-bit numbers (n1,n2,n3,n4) and the permutation function is obtained from its first excitation.

Permutation function							
4	1	2	3	2	3	4	1

As explained above, the comparison of the 4-bit values with permutation values, the following deliverables is given below

$$\begin{array}{c|c} n_4 & n_1 \\ n_2 & n_3 \end{array} \qquad \begin{array}{c|c} n_2 & n_3 \\ n_4 & n_1 \end{array}$$

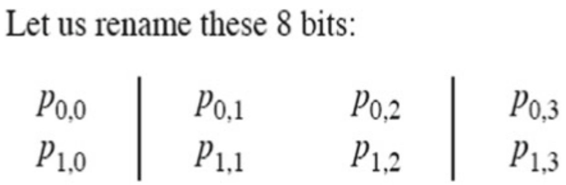
$K_{18=}$ 1000 1111.

K_{28} = 1000 1111.

IP = 1101 0101

$$\begin{array}{c|c} 0+1 & 1+0 \\ 0+1 & 1+1 \end{array} \qquad \begin{array}{c|c} 0+0 & 1+0 \\ 0+1 & 1+1 \end{array}$$

The XOR operations results in: 01,010,101.



Let us rename these 8 bits:

$$\begin{array}{c|cc|c} p_{0,0} & p_{0,1} & p_{0,2} & p_{0,3} \\ p_{1,0} & p_{1,1} & p_{1,2} & p_{1,3} \end{array}$$

10000 1111

(E) *DES Modes of Operations*

The encryption is DES algorithm converts the 64-bit message block M into 64-bit cipher block. If the encryption is processed individually for all the blocks, then this encryption mode is termed as electronic code book mode. Other than ECB mode, there are two other modes used in DES encryption such as cipher feedback (CFB) and chain block coding (CBC). In this modes, each cipher block is dependent on on all the previous messages blocks through an initial XOR operation.

(F) *Triple Data Encryption Standard (TDES)*

TDES is a symmetric key block cipher which is 3 times greater when multiplied with the DES cipher algorithm. The data encryption is performed with keys in a systematic order and TDES considers all the three 64-bit keys for an overall key length of 192 bits.

4 Input Design

The overall system design can be done with the help of input design, which requires a deliberate attention. Input data collection is an essential part of the system which depends on the equipment and number of people.

Making the data entry easy and free from logical error is the main objective of input data.

The objectives of the input design phase are to obtain a cost effective method to achieve the high accuracy level which makes the input as acceptable and understandable.

In this system, the table authentication used to check the user authorization. Then, we can choose tables to encrypt and decrypt.

5 Output Design

The changes made to input records, the normal procedure is to design the outputs in details first and then to work back to the inputs. The input records have to be validated, edited, organized, and accepted by the system before being processed to process to produce the outputs.

The result generated from the system is considered as output, and it is used to evaluate the utilization and effectiveness of the application. To obtain efficiency reports, software systems can be utilized. In this paper can layout the file register and file restore as a reports.

6 Implementation

System implementation phase is concerned with translation of the design specifications into the source code and internal documentation. This helps the testing, debugging, and modifications much simple, and the objective of the research can be obtained by projecting the source code in a clear manner. Generally, elegance and simplicity, clarity are the essential features of a good programs.

(A) *Maintenance*

The primary goal, of various phases of software engineering process discussed so far, is to improve the maintainability of software. That is, to improve the ease with which changes can be accommodated and to reduce the amount of time expanded on maintenance.

During the use of a large program, errors will occur and be reported to the developer. The error reported has to be tracked back to the source code (diagnosis) and corrected. The ease with which an error can be tracked back and corrected depends on how well the program is designed.

It is a software maintenance activity necessitated by change in the operating environment of the software. The environmental changes include, hardware changes, operating system changes. To accommodate these changes, the software has to be modified. Unlike corrective maintenance, this maintenance activity may require a complete design change in the software.

7 Conclusion

In this research work, the proposed algorithm was implemented by front end applications Asp.net, C#, and SQL server 2005 as a database. The data was fetched from database and stored in data table. This data was available in plain text format. This data was converted into bytes then converted into cyber bytes. These cyber bytes were converted into cyber text. These data were put into container and send to server through insecure channel. In other side, the data was received and decrypted, analyzed, and stored into database as single XML tag. The security was applied using encryption standards [Asymmetric and Symmetric Encryption]. In future, all symmetric and asymmetric encryption can be applied to the data as per user choice.

References

1. Nikita Gorasia RR, Srikanth ND, Rupareliya J (2016) Improving security in multi authority attribute based encryption with fast decryption. Procedia Comput Sci 79:632–639
2. Qian Q, Yu Z, Zhang R, Hung CL (2018) A multi-layer information dispersal based encryption algorithm and its application for access control. Sustain Comput Inf Syst 20:76–87
3. Bouhous A, Kemih K (2018) Novel encryption method based on optical time-delay chaotic system and a wavelet for data transmission. Opt Laser Technol 108:162–169
4. Sartakhti JS, Jalili S (2019) On the computational power of the light: a plan for breaking data encryption standard. Theoret Comput Sci 773:71–78
5. Masoumi M (2019) A highly efficient and secure hardware implementation of the advanced encryption standard. J Inf Secur Appl 48:1–12
6. Artiles JAP, Chaves DPB, Pimentel C (2019) Image encryption using block cipher and chaotic sequences. Signal Process Image Commun 79:24–31
7. Zhang K, Chen J, Lee HT, Wang H (2019) Efficient public key encryption with equality test in the standard model. Theoret Comput Sci 755:65–80
8. Crocett L, Baldanzi I, Bertolucci M, Fanucci L (2019) A simulated approach to evaluate side-channel attack countermeasures for the Advanced Encryption Standard. Integration 68:80–86
9. Hiscock T, Savry O, Goubin L (2019) Lightweight instruction-level encryption for embedded processors using stream ciphers. Microprocess Microsyst 64:43–52

10. Nardo LG, Nepomuceno EG, Arias-Garcia J, Butusov DN (2019) Image encryption using finite-precision error Chaos. Solitons & Fractals 123:69–78
11. Sussman D, Ullman C (2006) Beginning ASP.Net with C#, 2006 Edition
12. Nagel C, Skinner M (2004) Bill Evjen "Professional C#" Wiley publishing 2004, Edition
13. Mollin RA (2003) RSA and public-key cryptography, CRC Press 2003 Edition
14. Lewko A, Waters B (2010) Decentralizing attribute-based encryption, Cryptology ePrint Archive: Rep. 2010/351

Electrical Performance of Single Gate and Double Gate MOSFET to Optimize the Gate Length

James Lalthlamuana, Niladri Pratap Maity, L. L. K. Singh, and Reshmi Maity

Abstract This work investigated the electrical performances of a single gate (SG) and double gate (DG) metal oxide semiconductor field-effect transistors (MOSFETs) by varying the gate length. The electrical characteristics are analyzed and parameters like ON-state current, OFF-state current, and transconductance are considered. A similar study was conducted by changing the oxide material to high-k material (HfO_2). It is perceived that for gate length below 25 nm, overall channel potential starts decreasing due to increased source-drain resistance and velocity saturation owing to short channel effects. The drain current decreases as the gate length are decreased for both the SG and DG MOSFET. The transconductance also decreases inversely with the gate length for SG with the graded channel, SG with graded channel using a high-k dielectric, DG with the graded channel, and DG with graded channel using high-k dielectric.

Keyword Single gate (SG) and double gate (DG) metal oxide semiconductor field-effect transistors (MOSFETs)

1 Introduction

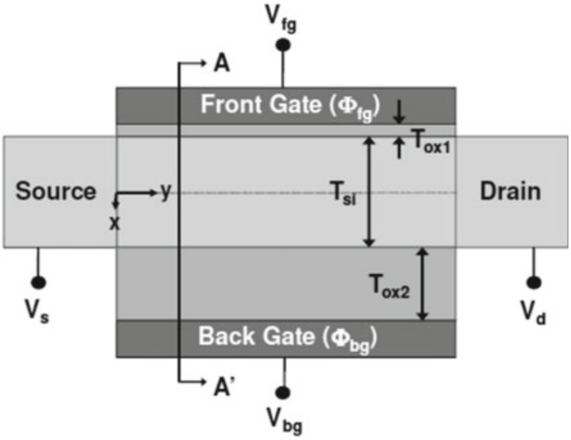
MOSFET has to turn into one of the most significant devices in the manufacturing industry because of its successful incorporation into integrated circuits (ICs). Even though the physical structure of MOSFET has not changed in full, its size has been repeatedly reduced by double every 2–3 years in accord with Moore's Law [1]. To look for probable substitutes for bulk MOSFETs beyond the 45 nm technology node, many innovative multi-gate MOSFETs have been suggested. They are gate-all-around (GAA) MOSFET [2], TFET [3], FinFET [4], Tri-Gate [5] and DG MOSFETs [6]. DG MOSFET is a freshly developing device that can more scale down complementary MOSFET (CMOS) technology in the sub-50 nm regime owing to its outstanding control of short channel effects.

J. Lalthlamuana · N. P. Maity (✉) · L. L. K. Singh · R. Maity
Department of Electronics and Communication Engineering, Mizoram University (A Central University), Aizawl 796 004, India
e-mail: maity_niladri@rediffmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_52

Fig. 1 Configuration of DG MOSFET



2 Configuration of DG MOSFET

Enhanced scalability must be talented by adding another gate terminal on the backside of the body terminal of the standard MOSFET ensuing in the double-gated configuration. The configuration of a double-gated FET is depicted in the inset to Fig. 1.

The downscaling of CMOS technology take on SiO_2 as insulator material has got hold of its maximum; there is an idea in diminishing the electrical oxide thickness more by using high-k materials [7–9]. Among the high-k applicants, HfO_2 is the greatest encouraging material [10–14].

This work focuses on the optimization of the electrical parameters in both SG and DG MOSFETs by the implementation of the graded channel and source/drain engineering and further, improving the performance of the device with the introduction of the high-k material. ON-state current (I_{ON}), OFF-state current (I_{OFF}), and transconductance (g_m) are considered for the analysis. Channel engineering is also employed to achieve the best performance with a graded channel which controls several short channel effects (SCEs) [15–18].

3 Results and Discussion

The electrical performances of SG and DG MOSFETs are analyzed and optimized by varying the gate length. The electrical characteristics are analyzed, and parameters like I_{ON}, I_{OFF}, and transconductance are considered. Polycrystalline silicon (poly-Si) is again used as gate material with work function set as 4.17 eV. Room temperature (300 K) is considered with ohmic contacts for source, drain, and substrate. Further analysis is done using the optimized device parameters. The proposed work has been implemented on different device structures using different dielectric materials and classified into: SG with a graded channel, SG with graded channel using a high-k dielectric, DG with a graded channel, and DG with a graded channel, and high-k dielectric. Figures 2 and 3 represent $I_d - V_g$ characteristic curves of SG MOSFET with the graded channel and SG MOSFET with graded channel using high-k dielectric, respectively. Figures 4 and 5 represent $I_d - V_g$ curves of the DG MOSFET with the graded channel and DG MOSFET with graded channel using high-k dielectric, respectively.

Figure 6 and 7 represent $I_d - V_d$ characteristic curves of SG MOSFET with graded channel and SG MOSFET with graded channel using high-k dielectric material, respectively. Figure 8 and 9 represent $I_d - V_d$ characteristic curves of DG MOSFET with graded channel and DG MOSFET with graded channel using high-k material, respectively.

Figures 10 and 11 represent $g_m - V_g$ characteristic curves of SG MOSFET with graded channel and SG MOSFET with graded channel using high-k dielectric material, respectively. Figures 12 and 13 represent $g_m - V_g$ characteristic curves of DG MOSFET with graded channel and DG MOSFET with graded channel using high-k material, respectively.

From the above graphs and Table 1, it is seen that I_{ON} is maximum at the gate length of 25 nm. For gate length below 25 nm, the overall channel potential starts decreasing due to increased source-drain resistance and velocity saturation owing to short channel effects. Equally seen, the drain current decreases as the gate lengths are decreased for both SG and DG MOSFETs. The transconductance also decreases inversely with the gate length for SG with graded channel, SG with graded channel using a high-k dielectric, DG with graded channel, and DG with graded channel using high-k dielectric. The transconductance has a peak value at the gate length of 25 nm.

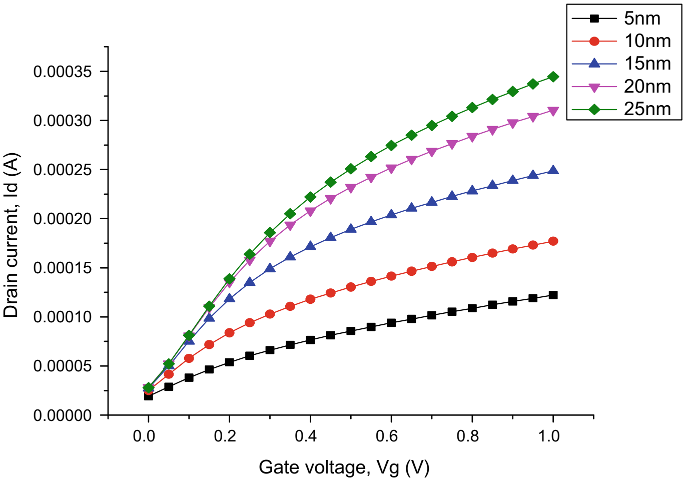


Fig. 2 $I_d - V_g$ plot of SG MOSFET

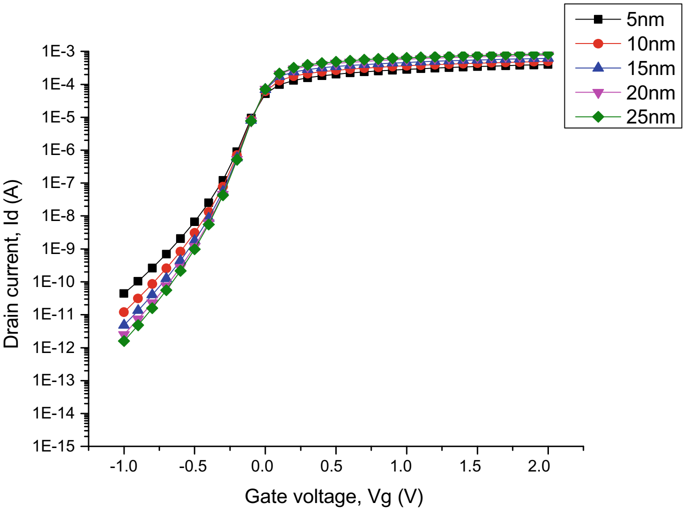


Fig. 3 $I_d - V_g$ plot of SG MOSFET with graded channel using HfO_2

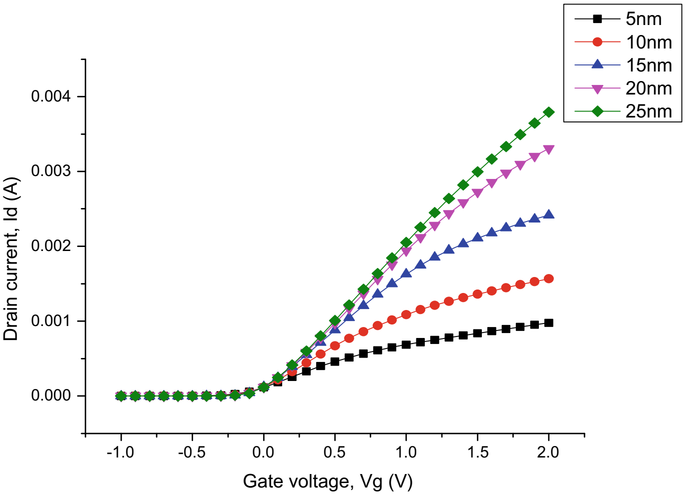


Fig. 4 $I_d - V_g$ plot of DG MOSFET

I_{OFF} should be ideally zero to have good subthreshold characteristics. The I_{OFF} is minimum when the gate length is 25 nm as the gate leakage current through the very thin dielectric material is minimum at this length. Hence, it can be said that the device performance is optimum for gate length, $L_G = 25$ nm for all the four cases.

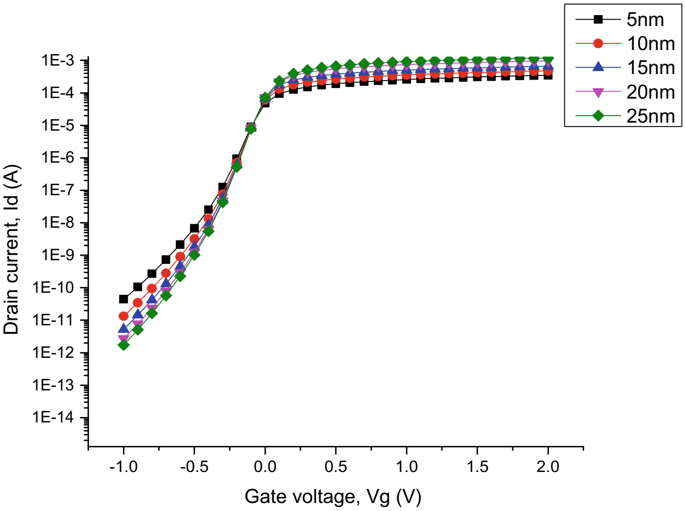


Fig. 5 $I_d - V_g$ plot of DG MOSFET with graded channel using HfO_2

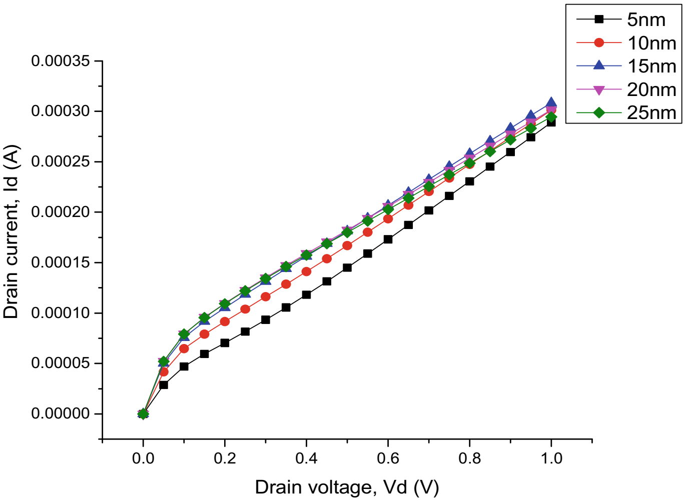


Fig. 6 $I_d - V_d$ plot of SG MOSFET

4 Conclusion

Single gate metal oxide semiconductor field-effect transistors have been optimized for better performance by using the graded channel and source/drain engineering with lightly doped drain structure to maintain satisfactory threshold voltage level, channel mobility, and punch through a mechanism. Channel engineering has been utilized

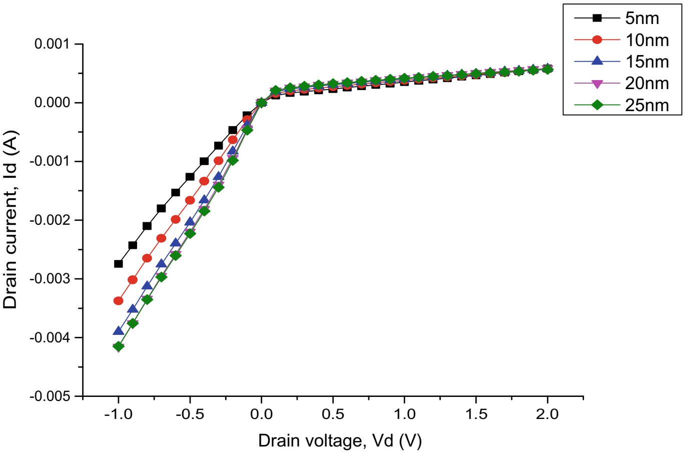


Fig. 7 $I_d - V_d$ plot of SG MOSFET with graded channel using HfO_2

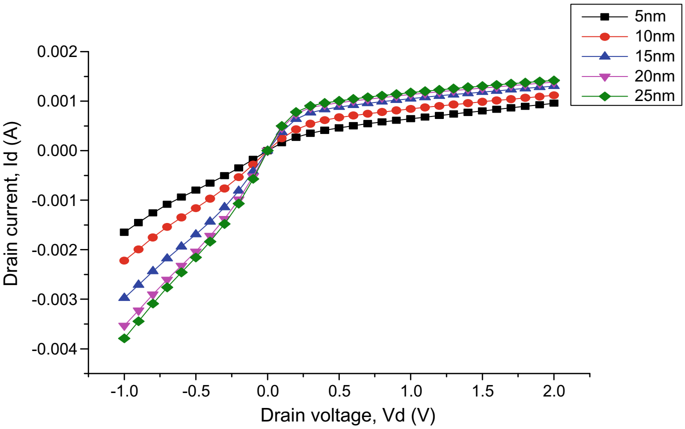


Fig. 8 $I_d - V_d$ plot of DG MOSFET

to obtain the best performance with the graded channel which controls the short channel effects. Then, the device was simulated for obtaining the best performance parameters by varying the gate length. A similar study was conducted by changing the device structure and material with changing the oxide material to HfO_2 (high-k material), the addition of second gate (DG MOSFET) and changing oxide material to HfO_2 in DG MOSFET. For all four cases, the gate length was different from 5 to 25 nm. The high current drive; i.e., peak drain current and peak transconductance values were observed for $L_G = 25$ nm. It was observed that DG has better performance than SG MOSFET in terms of current drive and transconductance. Also, the use of

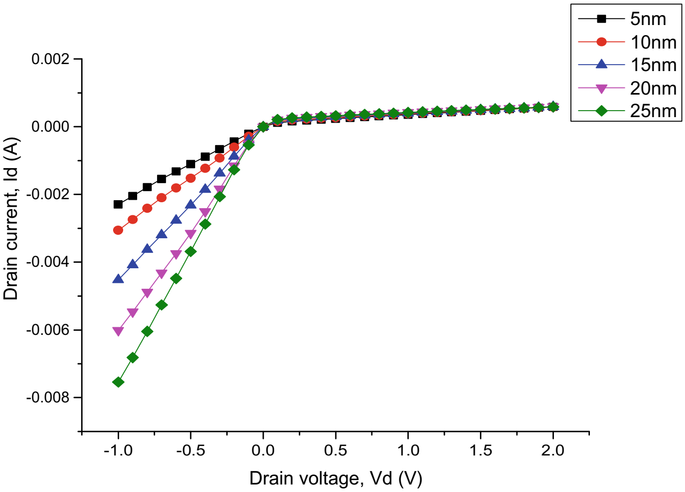


Fig. 9 $I_d - V_d$ plot of DG MOSFET with graded channel using HfO_2

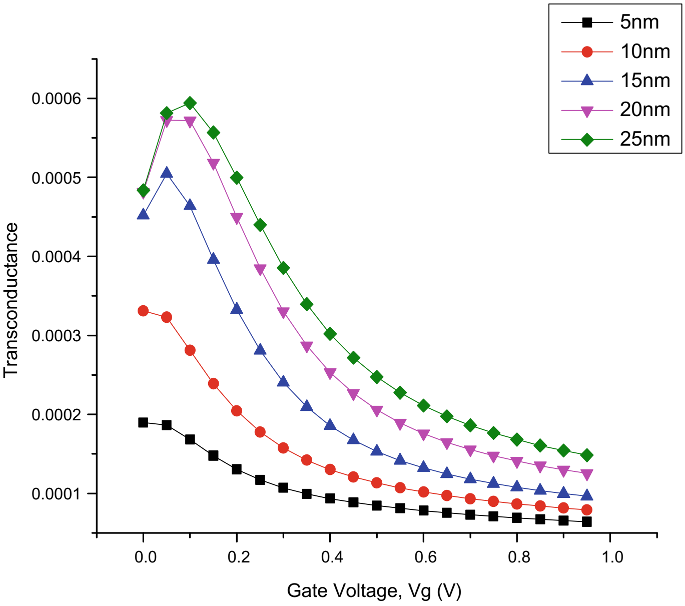


Fig. 10 $g_m - V_g$ plot of SG MOSFET

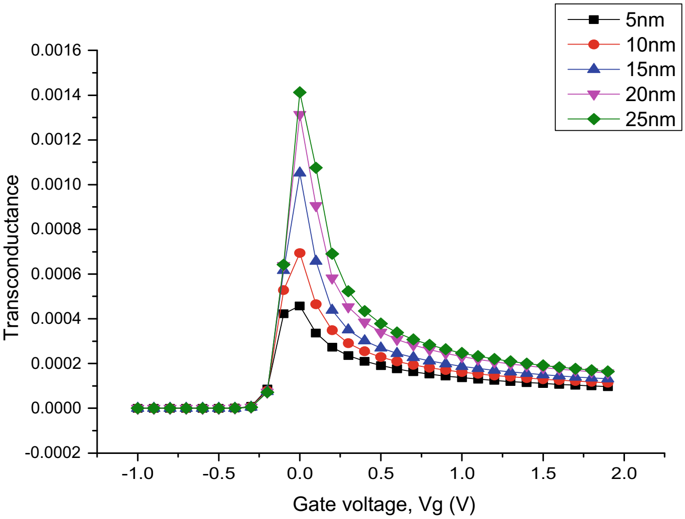


Fig. 11 $g_m - V_g$ plot of DG MOSFET with graded channel using HfO_2

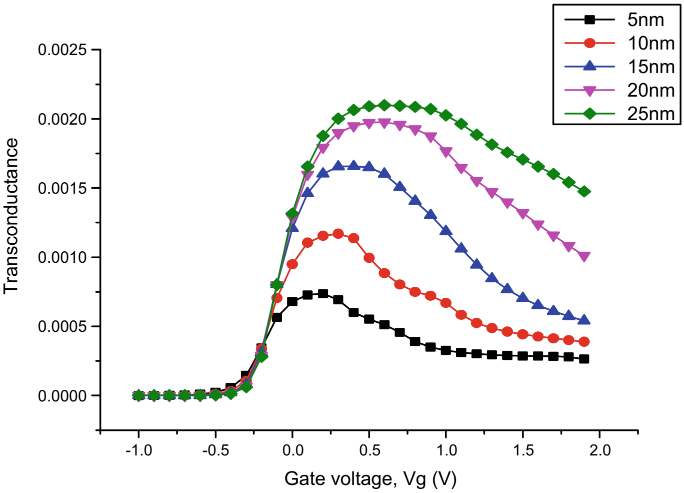


Fig. 12 $g_m - V_g$ plot of DG MOSFET

high-k material, i.e., HfO_2 in both single gate and double gate improves the electrical performance of the MOSFETs.

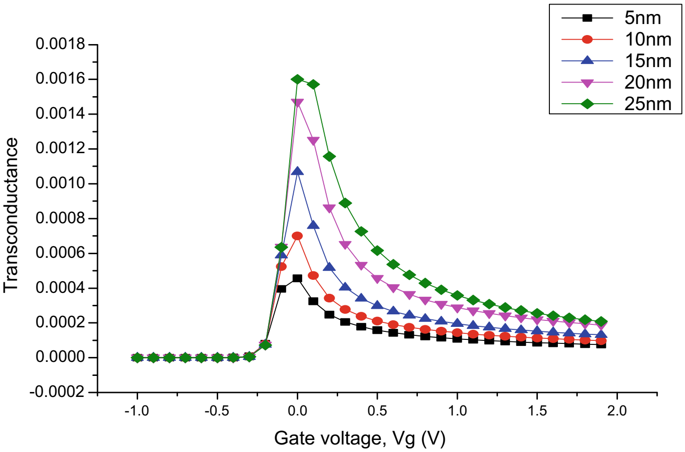


Fig. 13 $g_m - V_g$ plot of DG MOSFET with graded channel using HfO2

Table 1 Optimization of gate length

Description	Para-Meter	$L_G = 5$ nm	$L_G = 10$ nm	$L_G = 15$ nm	$L_G = 20$ nm	$L_G = 25$ nm
SG with graded channel	I_{ON}	0.00012	0.00017	0.00024	0.00031	0.00034
	I_{OFF}	1.919×10^{-5}	2.506×10^{-5}	2.748×10^{-5}	2.812×10^{-5}	2.8×10^{-5}
	g_{max}	6.408×10^{-5}	7.924×10^{-5}	9.642×10^{-5}	1.253×10^{-4}	1.4838×10^{-4}
SG with graded channel using HfO_2	I_{ON}	0.00040	0.00049	0.00062	0.00076	0.00084
	I_{OFF}	4.43×10^{-11}	1.21×10^{-11}	4.84×10^{-12}	2.5×10^{-12}	1.578×10^{-12}
	g_{max}	9.7×10^{-5}	0.0001141	0.0001314	0.0001581	0.00016442
DG with graded channel	I_{ON}	0.00097	0.00156	0.00241	0.00330	0.00379
	I_{OFF}	4.601×10^{-8}	9.332×10^{-9}	2.558×10^{-9}	1.107×10^{-9}	6.569×10^{-10}
	g_{max}	0.0002635	0.0003892	0.0005412	0.0010118	0.0014753
DG with graded channel using HfO_2	I_{ON}	0.00034	0.00047	0.00065	0.00096	0.00118
	I_{OFF}	4.44×10^{-11}	1.34×10^{-11}	5.292e-12	2.73×10^{-12}	1.73×10^{-12}
	g_{max}	7.58×10^{-5}	9.85×10^{-5}	0.0001317	0.000188	0.0002088

References

1. Moore G (1965) Cramming more components onto integrated circuits. Electron 38:114–117
2. Rana V, Ahmed G, Ramesh A, Das S, Singh P (2020) Diameter depended piezoresistive sensing performance of junctionless Gate-All-Around Nanowire FET. IEEE Trans Electron Dev 67(7):2884–2889
3. Li W, Woo J (2020) Vertical P-TFET with a P-Type SiGe Pocket. IEEE Trans Electron Devices 67(4):1480–1484

4. Maity NP, Maity R, Baishya S (2019) An analytical model for the surface potential and threshold voltage of a double-gate heterojunction tunnel FinFET. J Comput Electron 18:65–75
5. Ma J, Matioli E (2017) Slanted tri-gates for high-voltage GaN power devices. IEEE Electron Device Lett 38(9):1305–1308
6. Chakraborty H, Maity R, Maity NP (2019) Analysis of surface potential for dual-material-double-gate MOSFET based on modeling and simulation. Microsyst Technol 25:4675–4684
7. Maity NP, Maity R, Baishya S (2019) A tunneling current model with practical barrier for ultra thin high-k dielectric ZrO_2 material based MOS devices. Silicon 10:1645–1652
8. Maity NP, Maity R, Baishya S (2017) Influence of image force effect on tunneling current density for high-k material ZrO_2 ultra thin films based MOS devices. J Nanoelectron Optoelectron 12:67–71
9. Maity NP, Maity R, Thapa RK, Baishya S (2014) Study of interface charge Ddensities for ZrO_2 and HfO_2 based Metal-Oxide Semiconductor devices. Adv Mat Sci Eng Article ID 497274, 1–6
10. Chaudhry A, Kumar MJ (2004) Controlling short-channel effects in deep submicron SOI MOSFETs for improved reliability: A review. IEEE Trans Dev Mater Rel 4:99–109
11. Maity NP, Maity R, Baishya S (2017) Voltage and Oxide Thickness Dependent Tunneling Current Density and Tunnel Resistivity Model: Application to High-k Material HfO_2 Based MOS Devices. Superlattices Microstruct 111:628–641
12. Jelodar MS, Ilatikhameneh H, Kim S, Ng K, Klimeck G (2016) Optimum high-k oxide for the best performance of ultrascaled double-gate MOSFETs. IEEE Trans Nanotechnol 15:904–910
13. Choi J, Mao Y, Chang J (2011) Development of hafnium based high-k materials-a review. Mater Sci Eng R72:97–136
14. Chakraborty H, Maity R, Baishya S, Maity NP (2020) An accurate model for threshold voltage analysis of dual material double gate metal oxide semiconductor field effect transistor. Silicon, Online published in 9 July 2020https://doi.org/10.1007/s12633-020-00553-8
15. Basak R, Maiti B, Mallik A (2015) Analytical model of gate leakage current through bilayer oxide stack in advanced MOSFET. Supperlattices and Microstructures 80:20–31
16. Jelodar M, Ilatikhameneh H, Kim S, Ng K, Klimeck G (2016) Optimum high-k oxide for the best performance of ultrascaled double-gate MOSFETs. IEEE Trans Nanotechnol 15:904–910
17. Maity N, Maity R, Maity S, Baishya S (2019) Comparative analysis of the quantum FinFET and trigate FinFET based on modeling and simulation. J Comput Electron 18:492–499
18. Narang R, Saxena M, Gupta RS, Gupta M (2013) Impact of temperature variations on the device and circuit performance of tunnel FET: a simulation study. IEEE Trans. Nanotechnol. 12:951–957

Prediction of Diabetes Disease Using Machine Learning Model

Amandeep Sharma, Kalpna Guleria, and Nitin Goyal

Abstract As per the statistics mentioned by the world health organization, four hundred twenty-two million people in the world are suffering from diabetes which has raised the death toll to 1.6 million per year. This unprecedented growth in the number of cases and the number of casualties has led to an alarming situation because the data statistics represent a significant increase in diabetic cases among the young population, 18 years of age. Diabetes leads to various health hazards such as dysfunction of the kidney, cardiovascular problems, lower limb dismembering, and retinopathy. This article builds up a model for the prediction of diabetes using machine learning. The supervised machine learning algorithms used for prediction model such as decision tree, Naïve Bayes, artificial neural network, and logistic regression. Further, the comparison of these methods has been done based on various performance parameters such as accuracy, recall, precision, and F-score.

Keywords Machine Learning(ML) · Artificial neural network (ANN) · Logistic regression · Decision tree · Artificial intelligence (AI) · Naive bayes

1 Introduction

Diabetes is one of the common diseases nowadays. Earlier, the type 2 diabetes cases were reported primarily among the middle and old population. But in recent years, young children have also been reported as diabetic. The pancreas helps in the development of insulin in the human body. If the human body doesn't use the developed insulin efficiently and the pancreas doesn't develop essential quantities of insulin, then the human body gets diabetes [1]. Consequently, diabetes is treated as a major reason in global interest because of several health risks which may cause hyperglycemia [2]. High blood sugar is one of the major causes of diabetes, eye disease, heart attack, neuropathic ulcers [3, 4]. In [5], the authors have investigated worldwide diabetes prevalence and have found it more prominent in urban areas (10.8%)

A. Sharma · K. Guleria (✉) · N. Goyal
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
e-mail: kalpna@chitkara.edu.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_53

rather than rural (7.2%) areas. It has also been found that diabetes impacts more in high-income countries (10.4%) rather than countries which have low income (4.0%). Further, the authors have made a future forecast for diabetes worldwide and revealed that there would be 25% increased diabetic cases in 2030 and 51% diabetic cases in 2045. In the medical domain, it is an utmost important task to diagnose a disease at a primary stage because it will, in turn, lead to better treatment and slower progression of the disease or may prevent it further. Nowadays, machine learning plays an important role in the healthcare sector for the prediction of various diseases. So, it is extremely important to apply machine learning models for the initial detection of diabetes to improve the life of the human being. Modern technological improvement in the area of engineering and science relates to different artificial intelligence (AI) applications such as voice recognition, handwriting recognition, pattern matching, recommendation system, self-driving cars, stock market and real estate prediction, and medical healthcare. There are a large number of applications of artificial intelligence in the biocomputing field containing prediction of cardiac stroke, cancer classification, diabetic kidney disease assessment, and analytics [6–9]. In a machine learning process, the system learns from the previous experience and improves its performance [10–12]. It is a subset of AI which creates analytical models to build much statistical analytics.

2 Supervised Learning Algorithms

This section discusses various supervised learning classification algorithms which have been used for the construction of diabetic detection model. The dataset chosen for this work is Pima Indian Diabetic dataset, and it has been downloaded from UCI machine repository [13]. Below-mentioned supervised learning classification algorithms have been applied in this article:

(a) Decision Tree Algorithm
(b) Artificial Neural Network Algorithm
(c) Naive Bayes Algorithm
(d) Logistic Regression Algorithm

2.1 Decision Tree

Decision tree is an algorithm which comes under supervised learning techniques. To deal with continuous and categorical information, this method suits best [14, 15]. Decision trees do population splitting and splitting depends upon the splitter. The splitter can divide two or more than two same types of subsets. It forms a tree after completion of a task and its accuracy is majorly dependent on decisions. Classifying and applying regression on a tree follows dissimilar measures. One of the limitations of using DT is overfitting, which could be prevented by describing rules or restrictions on the building and pruning of a tree. Decision trees have numerous advantages in comparison to other supervised machine learning classification algorithms such as maximum likelihood classification. In addition, they tackle nonlinear relationships between features and classes, allow for disappeared values, and are capable of handling both categorical and numerical data inputs.

2.2 *Artificial Neural Network (ANN)*

The basic methodology of ANN is to function similarly to the human brain. ANN emulates the operation of the human brain. ANN architecture consists of various nodes known as artificial neurons which function similar to neurons in the human brain [16, 17]. It processes signals and then passes it further to the neurons connected to it. There exist various variants of ANN algorithms, however, the most popular ANN algorithm used is multilayer feed-forward ANN. It has input, output, and hidden layers. ANN also has limitations in finding the minima of global and the slow convergence degree. These layers are connected through many hidden nodes, and each node has a different weight. The root mean square error minimization between expected and predicted results is obtained by applying backpropagation using gradient descent optimization.

2.3 *Naive Bayes*

It is an approach, which uses the concept of Bayes theorem, by considering non-dependent interactive features in the designated dataset [18, 19]. The features of one class are always non-dependent on the features of another class. In the Naïve Bayes approach, all the assumptions are condition independent. This kind of algorithm could be utilized in the field of building a model where the dataset is having a huge amount of occurrences. Bayes theorem defined as below:

$$P(d/y) = (P(y/d)*P(d))/P(y)$$

.

where

P(d) signifies a class which is former for the P(x/c).

P(y/d) is used to denote likelihood probability.

P(d/y) identifies posterior probability.

P(y) denotes the probability of prior predictor.

It is also called an instant learner approach, which achieves the prediction conclusions very instantly for a class. It exhibits the best results for multiclass classification problems. In comparison with logistic regression, it outperforms as it needs less data for training. It has many applications, such as in-text identification, filtration of spam, recommender model, and sentimental analysis.

2.4 Logistic Regression (LR)

It is an approach which could be used for classification in machine learning model [20, 21]. There are many applications of LR including social studies and health sciences. It could also be used in the field of engineering, to identify and predict the probability of failure. For example, a water tank has two possible states, it could be safety or a failure, if there is an excess water supply, the binary LR will be used as the procedure to identify the relationship between safety and the predictor variable.

3 Predictive Model Construction for Diabetic Detection

The system model has been built with the help of labeled training dataset. The predictive model is required to understand the relationship between the input, output, and system variables [22]. In this paper, different supervised learning classifiers have been applied to construct a model for diabetic detection. The dataset chosen for this work is Pima Indian Diabetic dataset which has been collected from the UCI machine learning repository. This dataset contains data about females patients about 21 years of age and above. This labeled dataset has 768 instances and 9 attributes which include plasma glucose, diastolic blood pressure, triceps skinfold thickness, serum insulin, body mass index, diabetes pedigree, age, diabetes class variable. WEKA 3.8.4 [22, 23] simulator has been used to perform simulations.

Figure 1. shows different steps to create a predictive model for diabetic detection. The first step to developing any predictive machine learning model is a collection of raw data in a structured or unstructured form. After data collection, it is utmost important to perform the data cleaning to remove missing values and outliers. The next step is feature selection which selects various features from the given data set which are required and the most dominant one. Next, step is to train the model using machine learning model using a supervised learning classification algorithm. The percentage split up for training and testing used in this case is 70% and 30%, respectively. Thereafter, testing of the trained model is done using test data. The supervised learning model predicts results from its knowledge base of past experiences [24]. The performance of various classifiers is evaluated based on various performance parameters. The best classifier for this model is obtained based on accuracy value for the respective classifier, however, F-score is also one of the significant measures in this model since diabetic dataset represents an imbalanced class.

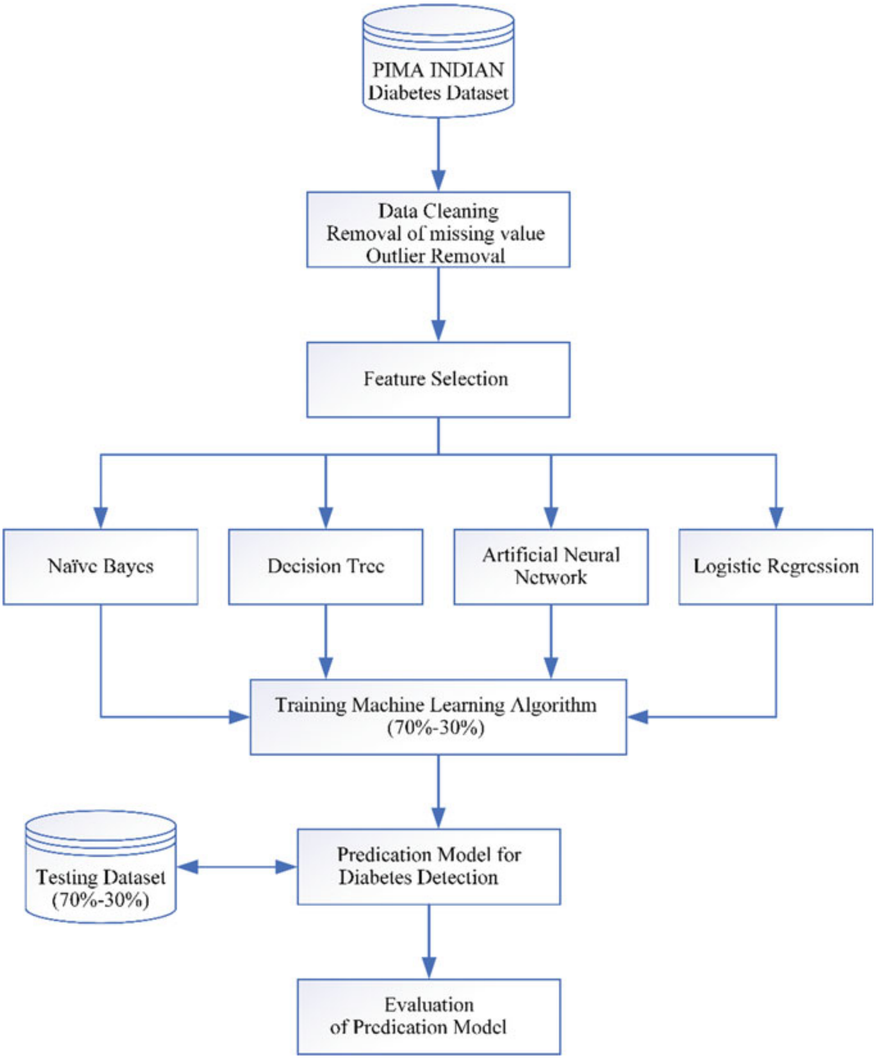

PIMA INDIAN
Diabetes Dataset
Data Cleaning
Removal of missing value
Outlier Removal
Feature Selection
Naïve Bayes
Decision Tree
Artificial Neural Network
Logistic Regression
Training Machine Learning Algorithm (70%-30%)
Testing Dataset (70%-30%)
Predication Model for Diabetes Detection
Evaluation of Predication Model

Fig. 1 Process of predictive model construction for diabetic detection

4 Results and Discussion

This section elaborates the results of prediction which has been obtained for various machine learning classifiers on this diabetic dataset.

Table 1. exhibits the simulation results of various classifiers on a diabetic dataset. Accuracy represents the number of right evaluation performed by the model. It identifies if the victim is diabetic or not. Recall (sensitivity) shows the segment of genuine diabetic patients [(tp/(tp + fn)]. Precision displays the percentage of positive diabetic

Table 1 Simulation results of various classifiers on a diabetic dataset

Classification	Naïve Bayes (%)	Decision tree (%)	ANN (%)	Logistic regression (%)
Precision	82.61	87.31	83.01	83.03
Recall	84.18	77.21	80.38	89.86
Accuracy	76.96	73.82	75.21	80.42
F-Score	83.39	81.87	81.66	86.31

cases, those are classified as diabetic or rightly identified [(tp/(tp + fp)]. It portrays the exact results or quality of the predictive model. Though, sensitivity exhibits quality or comprehensiveness of obtained result values. Large precision value shows that an algorithm gives more accurate results and higher sensitivity shows that most of the results are appropriate.

Table 2. shows the results of the Naïve Bayes algorithm where the total number of instances are 230, correctly classified instances are 177, and incorrectly classified instances are 53. The precision of this model is 82.61%, whereas recall for this model is 84.18%. F-measure for the given model is 83.39%. The overall accuracy of this model is 76.96%. The class-wise accuracy statistics about the true positive rate and false positive rate, along with F-score and ROC area has also been represented.

Table 3. represents the result of the decision tree algorithm. There are 176 instances correctly classified, and 54 instances are incorrectly classified. The precision of the given model is 87.13%, and recall for the algorithm is 77.21%. Accuracy of this algorithm is 73.82%. F-measure for the decision tree algorithm is 81.87%. Decision tree algorithms are used to predict the value of the target variable by understanding simple decision rules concluded from training data. Table 3 also presents class-wise accuracy statistics about the true positive rate and false positive rate, F-score, and ROC area.

Table 4 shows the result of ANN. In this classification algorithm, the correctly classified instances are 173 and incorrectly classified instances are 57. Precision for

Table 2 Simulation results of Naïve Bayes

Summary		
Correctly classified instances	177	76.9565%
Incorrectly classified instances	53	23.0435%
Mean absolute error	0.2677	–
Root mean squared error	0.3863	–
Total number of instances	230	–

Detailed Accuracy By Class

TP Rate	FP Rate	Precision	Recall	F-Score	ROC Area	Class
0.842	0.389	0.826	0.842	0.834	0.845	Tested_negative
0.611	0.158	0.638	0.611	0.624	0.845	Tested_positive
0.770	0.317	0.767	0.770	0.768	0.845	Weighted average

Table 3 Simulation results of decision tree

Summary		
Correctly classified instances	176	76.5217%
Incorrectly classified instances	54	23.4783%
Mean absolute error	0.3206	–
Root mean squared error	0.4239	–
Total number of instances	230	–

Detailed Accuracy By Class

TP Rate	FP Rate	Precision	Recall	F-Score	ROC Area	Class
0.772	0.250	0.871	0.772	0.819	0.743	tested_negative
0.750	0.228	0.600	0.750	0.667	0.743	tested_positive
0.765	0.243	0.786	0.765	0.771	0.743	Weighted average

Table 4 Simulation results of ANN

Summary		
Correctly classified instances	173	75.2174%
Incorrectly classified instances	57	24.7826%
Mean absolute error	0.3046	–
Root mean squared error	0.445	–
Total number of instances	230	–

Detailed Accuracy By Class

TP Rate	FP Rate	Precision	Recall	F-measure	ROC Area	Class
0.804	0.361	0.830	0.804	0.817	0.772	tested_negative
0.639	0.196	0.597	0.639	0.617	0.772	tested_positive
0.752	0.309	0.757	0.752	0.754	0.772	Weighted average

this model is 83.01%, and recall (sensitivity) for this model is 80.38%. Accuracy for the ANN algorithm is 75.21%. F-score of the model is 81.66%. The main feature of ANN algorithms is a non-parametric model; other than statistical models are parametric models and require a higher background of statistics. The class-wise accuracy statistics about the true positive rate and false positive rate, along with F-score and ROC area has also been represented.

Table 5. shows the results of logistic regression algorithm applied to Pima Indian diabetic dataset. Correctly classified instances are 185, and incorrectly classified instances are 45. In this classifier, the precision is 83.03% and recall for the applied dataset is 89.86%. The overall accuracy of the classifier is 80.42%. The F-measure score is 86.31% for the logistic regression algorithm. Logistic regression is better

Table 5 Simulation results of logistic regression

Summary						
Correctly classified instances		185		80.4348%		
Incorrectly classified instances		45		19.5652%		
Mean absolute error		0.2987		–		
Root mean squared error		0.3748		–		
Total number of instances		230		–		
Detailed Accuracy By Class						
TP Rate	FP Rate	Precision	Recall	F-measure	ROC Area	Class
0.899	0.403	0.830	0.899	0.863	0.848	tested_negative
0.597	0.101	0.729	0.597	0.656	0.848	tested_positive
0.804	0.308	0.799	0.804	0.799	0.848	Weighted average

than a decision tree because a single line is fit to divide space into two when the data is scattered in such a way so that it can be linearly classified. Table 3 also presents class-wise accuracy statistics about the true positive rate and false positive rate, F-score, and ROC area.

5 Conclusion

This research paper presents the prediction of a diabetic using machine learning models. supervised learning algorithms such as logistic regression, Naïve Bayes, artificial neural network, decision tree has been used to create the analytics models for finding whether the patient is diabetic or not. Accuracy represents the perfection of an algorithm. The prediction model exhibits that the logistic regression displays 80.43% accuracy which is highest among all. Naïve Bayes algorithm and decision tree display very competitive results. The accuracy of the Naïve Bayes algorithm is 76.95%, and decision tree algorithm has an accuracy of 76.52%; so, the final results of both classifiers are very close to each other. Artificial neural network classifier has 75.21% accuracy, which is the lowest among others. Further, rather than accuracy, F-score is also another effective measure to evaluate the prediction model. F-measure value can be represented between 0 to 1. If the F-measure value of any classifier is close to 1 means that the classifier model represents better performance. Logistic regression classifier represents 0.863 F-measure, which is highest among other classifiers and F-measure for the decision tree classifier is 0.817 lowest among other models. F- Measure for Naïve Bayes and ANN classifier is 0.834 and 0.819, respectively. Therefore, it is concluded that for this diabetic dataset, logistic regression represents the highest accuracy and F-score to create an analytical model for diabetes detection among other machine learning algorithms.

References

1. Kawada T (2020) Total dietary antioxidant capacity and risk of type 2 diabetes. Eur J Epidemiol 1–2. https://doi.org/10.1007/s10654-020-00608-5
2. Mamykina L, Heitkemper EM, Smaldone AM, Kukafka R, Cole-Lewis HJ, Davidson PG, Hripcsak G (2017) Personal discovery in diabetes self-management: discovering cause and effect using self-monitoring data. J Biomed Inform 76:1–8
3. Papatheodorou K, Banach M, Edmonds M, Papanas N, Papazoglou D (2015) Complications of diabetes.
4. Soumya D, Srilatha B (2011) Late stage complications of diabetes and insulin resistance. J Diabetes Metab 2(9):1000167
5. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, Shaw JE (2019) Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas. Diabetes Res Clin Pract 157:107843
6. Sapra L, Sandhu JK, Goyal N (2020) Intelligent method for detection of coronary artery disease with ensemble approach. Adv Commun Comput Technol (pp 1033–1042). Springer, Singapore
7. Guleria K, Sharma A, Lilhore UK, Prasad D (2020) Breast Cancer prediction and classification using supervised learning techniques. J Comput Theor Nanosci 17(6):2519–2522
8. Lilhore UK, Simaiya S, Prasad D, Guleria K (2020) A Hybrid Tumour detection and classification based on machine learning. J Comput Theor Nanosci 17(6):2539–2544
9. Babič F, Majnarić L, Lukáčová A, Paralič J, Holzinger A (2014) On patient's characteristics extraction for metabolic syndrome diagnosis: predictive modelling based on machine learning. Int Conf Inf Technol Bio Med Inf (pp 118–132) Springer, Cham
10. Chlingaryan A, Sukkarieh S, Whelan B (2018) Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review. Comput Electron Agric 151:61–69
11. Goyal N, Sandhu JK, Verma L (2019) Machine learning based data agglomeration in underwater wireless sensor networks. Int J Manage Technol Eng 9(6):240–245
12. Lilhore UK, Simaiya S, Guleria K, Prasad D (2020) An efficient load balancing method by using Machine Learning-based VM distribution and dynamic resource mapping. J Comput Theor Nanosci 17(6):2545–2551
13. Bay SD, Kibler D, Pazzani MJ, Smyth P (2000) The UCI KDD archive of large data sets for data mining research and experimentation. ACM SIGKDD Explorations Newsl 2(2):81–85
14. Ming H, Wenying N, Xu L (2009) An improved decision tree classification algorithm based on ID3 and the application in score analysis. In 2009 Chinese control and decision conference (pp 1876–1879) IEEE
15. Dong X, Qian M, Jiang R (2020) Packet classification based on the decision tree with information entropy. the Journal of Supercomputing 76(6):4117–4131
16. Momeni E, Nazir R, Armaghani DJ, Maizir H (2014) Prediction of pile bearing capacity using a hybrid genetic algorithm-based ANN. Measurement 57:122–131
17. Koopialipoor M, Fahimifar A, Ghaleini EN, Momenzadeh M, Armaghani DJ (2020) Development of a new hybrid ANN for solving a geotechnical problem related to tunnel boring machine performance. Eng Comput 36(1):345–357
18. Loor M, De Tré G (2020) Contextualizing Naive Bayes predictions. In International conference on information processing and management of uncertainty in knowledge-based systems (pp 814–827). Springer, Cham
19. Zhang H, Jiang L, Yu L (2020) Class-specific attribute value weighting for naive bayes. Inf Sci 508:260–274
20. Yang Y, Chen G, Reniers G (2020) Vulnerability assessment of atmospheric storage tanks to floods based on logistic regression. Reliab Eng Syst Saf 196:106721
21. Shah K, Patel H, Sanghvi D, Shah M (2020) A comparative analysis of logistic regression, random Forest and KNN models for the text classification. Augmented Human Research 5(1):1–16

22. Tiwari S, Kumar S, Guleria K (2020) Outbreak trends of Coronavirus Disease–2019 in India: a prediction. Disaster medicine and public health preparedness, pp 1–6
23. Frank E, Hall M, Holmes G, Kirkby R, Pfahringer B, Witten IH, Trigg L (2009) Weka-a machine learning workbench for data mining. In Data mining and knowledge discovery handbook (pp 1269–1277). Springer, Boston, MA
24. Singh A, Thakur N, Sharma A (2016) A review of supervised machine learning algorithms. In 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom) (pp 1310–1315) IEEE

Infrared Small Target Detection Based on Phase Fourier Spectrum Transform

Sur Singh Rawat, Sashi Kant Verma, and Yatindra Kumar

Abstract Recently, target detection task in infrared imaging is becoming very difficult when the high detection rate as well as low error rate is concerned. As the image background is noisy and cluttered, present methods face issues with detection performance and low error rate. To improve the detection ability of the existing methods, human visual saliency based method is presented in this letter. Firstly, it detects the salient region in the infrared image, which may contain the targets with the help of phase spectrum Fourier transform (PET). Then, saliency map is obtained and at the end simple threshold operation is performed to get the candidate target. Lot of experiments were conducted and the results show that the method presented here can lower the background noise effectively and also predicts the object efficiently with reduced error rate.

Keywords Infrared small target detection · Saliency · Phase spectrum · Fourier transform · Human visual system

1 Introduction

Infrared technology finds a large number of applications in defense and military system. Target detection is found of the well-known task in many areas like space system, warning systems as well as in missile system. The target which is small is normally get remained in a noisy background along with less signal to noise ratio and there is always a long imaging distance and atmospheric transmission.

S. S. Rawat (✉)
Department of Computer Science and Engineering, JSS Academy of Technical Education, Noida, U.P., India
e-mail: sur.rawat@jssaten.ac.in

S. K. Verma
Department of Computer Science and Engineering, GBPIET, Pauri garhwal, Uttrakhand, India
e-mail: skverma.gbpec@rediffmail.com

Y. Kumar
Department of Electrical Engineering, GBPIET, Pauri garhwal, Uttrakhand, India
e-mail: kyatindra@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_54

Although many research work is done in this direction in the past decades but the above said circumstances make the target detection task still a hard problem [1–3]. To tackle with the above said issues, many methods have been addressed in the past. These methods are basically broken into two folds: Single frame methods and multiple frames sequentially methods. The second methods make full use of spatial and temporal features acquired by imaging sensor, but, as the detected target or the sensor system is of high speed in many applications, so this produces some kind of effects due to the effective difference in the image with the background and the target. As a consequence of this, the sequential method does not work well. As the detection of target required detection performance to be high with reduced response time, then a single frame method is a better choice. Gao et al.tried a method to detection of target issue and come-up with a method, patch Image-model (IPI) [1].This method has utilized, non- local self-correlation features of background. Also, expect the background patches are normally belongs to the combination or the single low-rank subspace clusters. The IPI model has the l_1-norm sparsity problem as well as the constant weighting parameter issue which estimates the background inaccurately and this has produced a miss classified target object. Dai et al. [4] Gave a method which was on structural properties of the images in background. This approach although has shown better result when compared with the other approaches, but here in this method, an extra load was created in computation of weights of the columns. Dai et al. [5] presented a non-negative IPI method based on partial sum of minimization of singular values (PSMSV) (NIPPS) so as to approximate the background and retained the singular values. NIPPS has a problem of choosing a proper energy constraint ratio as well as the rank of the matrix. Gao et al. [6], Reweighted IPI (ReWIPI) model was based on work discussed in [7] to adjust the patch image of background and also retain the background edge features. Problem with this approach was the improper tuning of weighting parameter which could affect the calculation of singular value decomposition (SVD) of matrix. To, further improve the IPI based approaches, Non-convex rank approximation minimization (NRAM) [8] was presented. Also, to bring the inner smoothness to the background, a term for regularization called total variation was applied to the method namely: regularization of TV and induction of principal component pursuit (TV-PCP) [9]

Human visual system (HVS) [10] was presented recently to target detection system in infrared imaging, which think that the target is one of the salient information. If we consider methods based on saliency, Laplacian of Gaussian (LOG) [11, 12],Difference of Gaussian (DOG) [13],second-order directional derivative (SODD) [14],contrast measure in local area (LCM) [15],patch-based method on contrast measure and multi-patch image(MPCM) [16], method on difference of Gabor’s an improved version (IDOGb) [17],difference measure in local area(LDM) [18],weighted local difference measure(WLDM) [19], spectral residual [20], phase spectrum [21] are the most popular one.

The remaining paper is presented as follows: Section 2 here includes the description the related work, Sect. 3, we describe the process target detection for small objects. In Sect. 4, include the result of method presented and its comparison with the other methods. Section 5 presents the concluding of this paper.

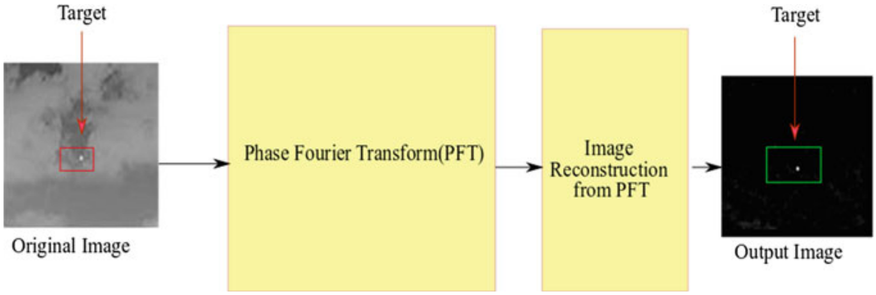


Fig. 1 Proposed method for small target detection

2 Methodology: Method Based on Human Visual System

2.1 *Brief Description of a Human Visual System*

(HVS) basically break the input image into the different small patches and select the salient features through HSV system to do the easy understanding and analyze. We can observe in the HVS system that, important features in the image can be represented not by the amplitude of the visual signal, but it is actually the contrast information with the amplitude at a given point and at its vicinity location. From this, it is concluded that contrast is the valuable feature which is presented in streams of our visual system. Based on the above discussion we proposed a framework for target detection system as shown in (Fig. 1). HVS process will help to separate the target from background noise. Since the background noise are lke infrared target shape and size, so it is very difficult to distinguish them using the traditional methods.

2.2 *Theory of Phase Fourier Transform*

The phase Fourier transform (PFT) [20]of an image signal gives the phase information in an image and tells the location of an object present in an image. Phase spectrum information in an infrared image will detect the target as it the most salient information.

Following steps are needed to calculate the phase Fourier transform of an image. Let $s(x,y)$ is an image

$$f(x, y) = F(s(x, y)) \tag{1}$$

$$p(x, y) = P(f(x, y)) \tag{2}$$

$$SM(x, y) = g(x, y) * \|F^{-1}\big[e^{i.p(x,y)}\big]\|^2 \tag{3}$$

Where F and F^{-1} are the Fourier and its inverse Fourier transform,$P(f(.))$ is the phase spectrum of an image and $g(x, y)$ is the 2-D Gaussian filter at ($\sigma = 8$). $SM(x, y)$ is the saliency map of an image.

2.3 Steps of the Detection Method for Small Object

PFT is used to locate the salient object in an image so in our problem small object is the salient object that we need to detect. Following are the steps to perform in the proposed method.

1. PFT of the input infrared image is computed.
2. Saliency map is obtained after inverse PFT.

Finally the simple threshold operation is adopted to get the object more accurately.

3 Experimental Result and Analysis

We have prepared a dataset of more than 450 images under infrared imaging system of a varying environment such as, sea, sky, cloud and ground. Description of these images sequences are given in the Table 1. All experiments were implemented on MATLAB 2015 software on PC with 4-GB RAM and 2.20 GHz Intel Core 2 Duo processor.

3.1 Suppression Result of Image Sequences for Background

In this sub-section we have presented the background suppression result of the method presented of six different image sequences under the complex background. Figure 2a represents the original infrared images, Fig. 2b shows the background suppression result of our method. The location of target in an image can be seen by utilizing 3D gray map and finally, to show the background suppression result of our proposed method as shown in the Fig. 2c. The proposed approach is also compared to the baseline method shows the high background suppression ability and the high signal to clutter ratio gain. Figure 3 shows the effectiveness of our method with the other baseline method. As shown in the Fig. 3b, c the Max-mean and the max-median methods can predict the object well but in the strong clutter background these methods fail to detect the target effectively. Similarly Top- hat method as shown in Fig. 3d need to adjust the size of the filter to detect the target. Again in the strong clutter background it sometimes fail to detect the target. The

Table 1 Detail description of representative images

Real infrared image sequences	Target type	No of frames	Image size	Features of background	Features of target
Sequence no. 1	A small ship	30	256 × 200	Sea-sky with blurred	Changing target size • Target in small size • Imaging distance is long
Sequence no.2	An airplane	250	256 × 200	Heavy cloudy background containing low local contrast	Changing target size • Target in small size • Imaging distance is long
Sequence no. 3	An airplane	250	256 × 200	With changing background	Changing target size Target in small size Imaging distance is long Low SCR
Sequence no. 4	A Helicopter	100	241 × 200	Changing background	Imaging distance is long • Low SCR
Sequence no. 5	A ship	100	320 × 240	Changing background	• A one to two target • Small in size
Sequence no. 6	A person, moving back and forth in forest	250	280 × 228	Background containing heavy clouds	Changing target size, • Imaging distance is long

method presented compute the phase Fourier transform of the image and detect the location of a salient information. As infrared small target is the salient object in the image so its location can be detected easily by the method presented. Figure 3e shows the detection result.

3.2 Evaluation Indicators

In order to judge our method, the two important classical evaluation metrics namely: signal to clutter ratio gain (SCRG) and background suppression factor (BSF) are

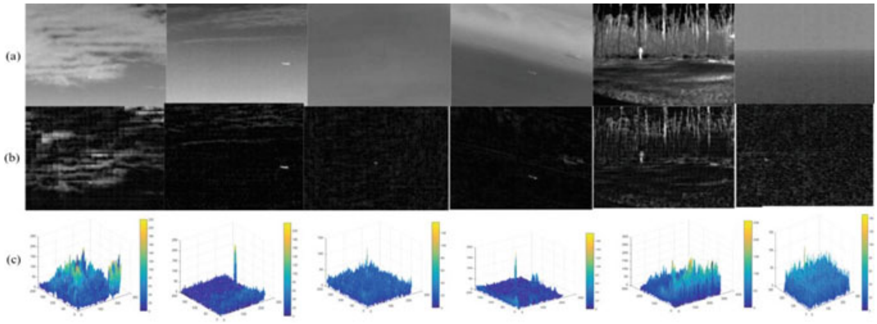


Fig. 2 3-D map of original infrared images and the result of the presented method. **a** Original representative images **b** Detection result of **a**, **c** 3-D map of proposed method respectably

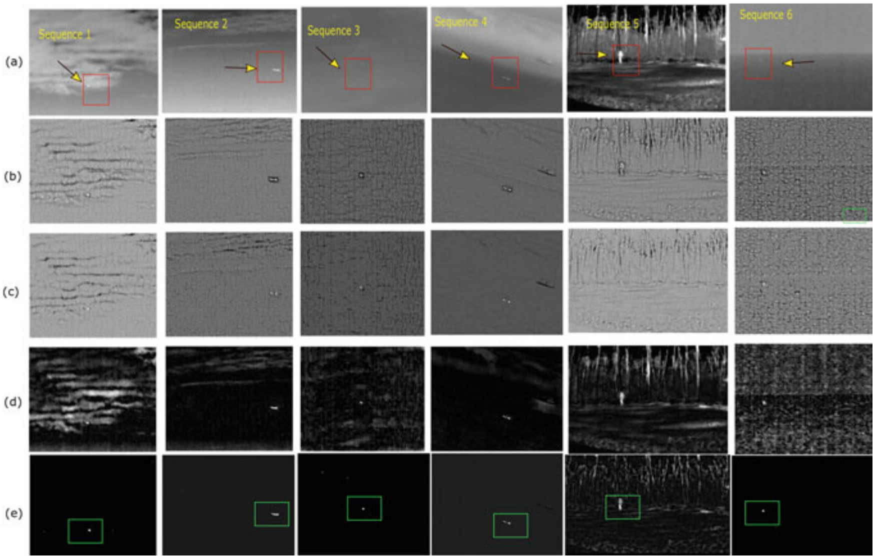


Fig. 3 Result of various methods **a** Input image **b** Max-mean **c** Max-median **d** Top-hat and **e** result of proposed method

utilized to validate the performance of the method. The detail of these metrics is given in [22] and can be written as follows:

$$\mathrm{SCRG} = \frac{\left({}^{S}/_{C}\right)_{\mathrm{out}}}{\left({}^{S}/_{C}\right)_{\mathrm{in}}}, \qquad \mathrm{BSF} = \frac{C_{\mathrm{in}}}{C_{\mathrm{out}}} \tag{4}$$

Here, signal of the amplitude and the clutter standard deviation is presented by S and C, *in and out* in the expression are input real image and the output image with target. The signal to clutter ratio gain (SCRG) tells the amplification result of the signal before and after the image is processed. Background suppression factor (BSF) given the level of suppression when no information of target is available. Hence it is expected that, for better efficiency both the indicators should have a large value.

3.3 Experimental Result Comparison with the Other Methods

The method presented here in this paper was compared with the **three (03)** other base methods to evaluate its robustness. Figure 3a depicts the original image sequences. The background suppression results of all the base methods like top-hat, max-mean and max-median along with the proposed method are presented in the Fig. 3b–e . We can observe that due to fractional directional derivatives and the phase Fourier transform the presented method lower down the background and improves the target better than the other base method where a simple filtering approach is performed. We can observe from Tables 2 and 3 which represents the signal to clutter ratio gain (SCRG) and the background suppression factor (BSF) respectably, that the method presented, has got better SCRG and BSF result in comparison with the other base methods. Similarly, as the time complexity is also an important parameter in an algorithm. The proposed method has computational efficiency of **0.051 s** which is better than the baseline methods as can be observe from Table 4.

Table 2 SCRG values of the presented method

Methods	Image sequences					
	A	B	C	D	E	F
Max-mean	1.16	5.73	7.38	12.74	1.32	12.13
Max-median	1.06	1.71	5.41	1.59	0.26	6.75
Top-hat	0.99	1.58	6.46	13.06	0.09	6.92
Proposed method	2.50	10.30	15.85	35.26	4.16	25.50

Table 3 Values of BSF of the presented method

Methods	Image sequences					
	A	B	C	D	E	F
Max-mean	0.49	2.34	0.51	3.93	1.34	0.92
Max-median	1.29	3.90	0.75	7.25	2.12	1.20
Top-hat	1.38	3.39	0.86	10.11	2.43	1.26
Proposed method	2.88	3.96	0.89	14.04	3.82	2.12

Table 4 Computation time of the method presented and base method

Methods	Times(s)
M ax-mean	0.841
Max- median	0.914
Top -hat	0.715
Proposed method	0.0583

4 Conclusion

An approach for target using the infrared images, which is motivated by the human visual system is presented in this paper. The phase Fourier transform is first used to obtain a saliency map of the candidate target and then a simple threshold approach is applied to get the resultant target. The experimental result shows that the method will minimize the noise from the background properly and also detect the candidate object effectively with the lesss error rate.

Acknowledgements The authors would like to thank all the valuable reviewers

References

1. Gao C, Meng D, Yang Y, Wang Y, Zhou X, Hauptmann AG (2013) Infrared patch-image model for small target detection in a single image. IEEE Trans Image Process 22(12):4996–5009
2. Zhao J, Tang Z, Yang J, Liu E (2011) Infrared small target detection using sparse representation. Journal of Systems Engineering and Electronics 22(6):897–904
3. Rawat SS, Verma SK, Kumar Y (2020) Review on recent development in infrared small target detection algorithms. Gurgaon, India
4. Dai Y, Wu Y, Song Y (2016) Infrared small target and background separation via column-wise weighted robust principal component analysis. Infrared Phys Technol 77:421–430
5. Dai Y, Wu Y, Song Y, Gao J (2017) Non-negative infrared patch-image model: Robust target-background separation via partial sum minimization of singular values. {Infrared Physics & Technology, vol 81, pp 182--194
6. Guo J, Wu Y, Dai Y (2017) Small target detection based on reweighted infrared patch-image model. IET Image Proc 12(1):70–79
7. Gu S, Xie Q, Meng D, Zuo W, Feng X, Zhang L (2017) Weighted nuclear norm minimization and its applications to low level vision. Int J Comput Vision 121(2):183–208
8. Zhang L, Peng L, Zhang T, Cao S, Peng Z (2018) Infrared small target detection via non-convex rank approximation minimization joint l2, 1 norm. Remote Sensing 10(11):1–34
9. Wang X, Zhenming P, Dehui K, Zhang P, He Y (2017) Infrared dim target detection based on total variation regularization and principal component pursuit. Image Vision Comput 63: 1--9
10. Itti L, Koch C, Niebur E (1998) A model of saliency-based visual attention for rapid scene analysis. IEEE Trans Pattern Anal Mach Intell 20(11):1254–1259
11. Kim S (2011) Min-local-LoG filter for detecting small targets in cluttered background. Electron Lett 47(2):105–106
12. Kim S, Yang Y, Lee J, Park Y (2009) Small target detection utilizing robust methods of the human visual system for IRST 30(9):994--1011

13. Wang X, Lv G, Xu L (2012) Infrared dim target detection based on visual attention. Infrared Physics & Technology}, vol. 55, no. 6, pp. 513--521, 2012.
14. Qi S, Ma J, Tao C, Yang C, Tian J (2013) A robust directional saliency-based method for infrared small-target detection under various complex backgrounds. IEEE Geosci Remote Sens Lett 10(3):495–499
15. Chen CP, Li H, Wei Y, Xia T, Tang YY (2014) A local contrast method for small infrared target detection. IEEE Trans Geosci Remote Sens 52(1):574–581
16. Wei Y, You X, Li H (2016) Multiscale patch-based contrast measure for small infrared target detection. Pattern Recogn 58:216–226
17. Han J, Ma Y, Huang J, Mei X, Ma J (2016) An infrared small target detecting algorithm based on human visual system. IEEE Geosci Remote Sens Lett 13(3):452–456
18. Deng H, Sun X, Liu M, Ye C, Zhou X (2017) Entropy-based window selection for detecting dim and small infrared targets. Pattern Recognition 61:66--77
19. Deng H, Sun X, Liu M, Ye C, Zhou X (2016) Small infrared target detection based on weighted local difference measure. IEEE Trans Geosci Remote Sens 54(7):4204–4214
20. Hou X, Zhang L (2007) Saliency detection: a spectral residual approach. In: IEEE conference on computer vision and pattern recognition IEEE, pp 1--8
21. Guo C, Ma Q, Zhang L (2008) Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform. In: 2008 IEEE conference on computer vision and pattern recognition, IEEE, pp 1--8.
22. Hilliard CI (2000) Selection of a clutter rejection algorithm for real-time target detection from an airborne platform. In: Proceedings SPIE, vol 4048, pp 74--84

Knowledge Based Analytical Tool for Identifying Children with Dyscalculia

A. Devi and G. Kavya

Abstract Specific Learning Disability is a learning difficulty that has an impact on reading (Dyslexia), mathematical calculations (Dyscalculia), and drawing skill (Dysgraphia), among children. Out of these learning disabilities, the mathematical disability is a major problem for children to do minor calculations like counting numbers, remembering phone numbers and problem in understanding speed, time, distance etc., which also affects their regular academic activities. Dyscalculia is a condition that lacks numerical skill needed for an individual that arises due to heredity or sometimes the brain development itself. Mostly children's who are affected with dyscalculia seem to be good in other learning areas. Due to the lack of identification and assessment, these children severely affected when they are entering into higher grades. This may weaken their mental health condition and make them feel left out and low in confidence. Hence a knowledge-based tool is proposed to identify mathematical difficulties using machine learning decision tree algorithm. The entire proposed model is divided into two sections (i) Knowledge-based tool creation for dataset collection (ii) Decision tree classification. The main aim of this research work is to diagnose and detect the dyscalculia problem in the children at the earlier stage itself.

Keywords Decision tree (DT) · Machine learning (ML) · Random forest (RF) · Content management system (CMS)

1 Introduction

Learning Disability (LD) is a neurodevelopmental disability that affects brain functioning capacity to process the listening information, observing information, etc. [1].

A. Devi (✉)
Department of ECE, IFET College of Engineering, Villupuram, Tamilnadu, India
e-mail: deviarumugam02@gmail.com

G. Kavya
Department of ECE, SA Engineering College, Chennai, Tamilnadu, India
e-mail: kavyavimal@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_55

The learning disability maybe appears in a reading, writing and mathematical skills, hence it is called Specific Learning Disability (SLD). Around worldwide, 10 percent of the students are affected by SLD. Statistics in the west and India show that about 13 to 14% of the children in a regular classroom have SLD. Specifically, around 6 percent of school-going children have math deficits. Such mathematical deficits were common and required equal attention and concern [2]. Effect of math failure with mathematical analphabetism seriously impedes both everyday life and studies. Many factors cause dyscalculia problem in the children. And also several assessment tests are available to screen the children with dyscalculia. Each assessment test focus on the different skills of the children. The relevant assessments can be assessed to [3]

(a) Computation skills—E.g. Woodcock-Johnson IV (WJIV),
(b) Math fluency—E.g. WJ IV Math Fluency, Wechsler Individual Achievement Test (WIAT-III) Math Fluency subtest, Mathematical Fluency and Calculations Tests (MFaCTs)
(c) Mental computation—E.g. Wechsler Intelligence Scale for Children (WISC-V), Paced Auditory Serial Addition Test (PASAT), and Test of Mental Computation
(d) Quantitative reasoning—WIAT-III Math Problem Solving subtest, WJ IV Applied Problems, CMAT Problem Solving.

These assessment tests can help to determine the dyscalculia issue of the children manually. At present Machine Learning is [4] widely used in many fields to predict future results. One of the most important applications of machine learning is to predict the learning disorder problem in children, identify the exact disability and recognise the disability problem at the earlier stage. The fundamental concept behind machine learning is to develop an algorithm that can accept input data and use statistical models to predict an exact output. Classification in machine learning consists of two steps such as learning and the prediction. In the learning step [5], the model is built based on provided training data. The model is used in the prediction process to predict the response to the data provided. For understanding and analysing the datasets, the commonly used classification algorithm is the decision tree algorithm, where the decision tree belongs to supervised learning algorithms. The main objective of the proposed tool is to accurately identify the children with dyscalculia problem at the earlier stage using machine learning Decision Tree algorithm.

2 Related Work

Let us discuss some of the studies carried out to identify the children with dyscalculia problem. And also screening tools developed for diagnosing and identifying the dyscalculia problem is discussed.

3 Proposed Knowledge-Based Tool

E-platform for dataset generation is created for identifying children with dyscalculia, where it consists of many questions which will be plugged in word press Content Management System (CMS) software platform by using the plugin method. The word press is an open-source platform [11], so that website can be customized, updated and managed by the user. Steps for the plugin is given below.

Step 1: Download Word press and log in to word press by using username and password which given during installation.

Step 2: Add details into general settings such as site title, website, email address, time format, date format etc.

Step 3: Install and configure a suitable plugin. In the proposed model, the Quiz Maker plugin and the Audio plugin is configured.

Once the plugin is installed, the questionnaire is set to meet the WJ IV test. The tests are conducted in quiz type (Fig. 1).

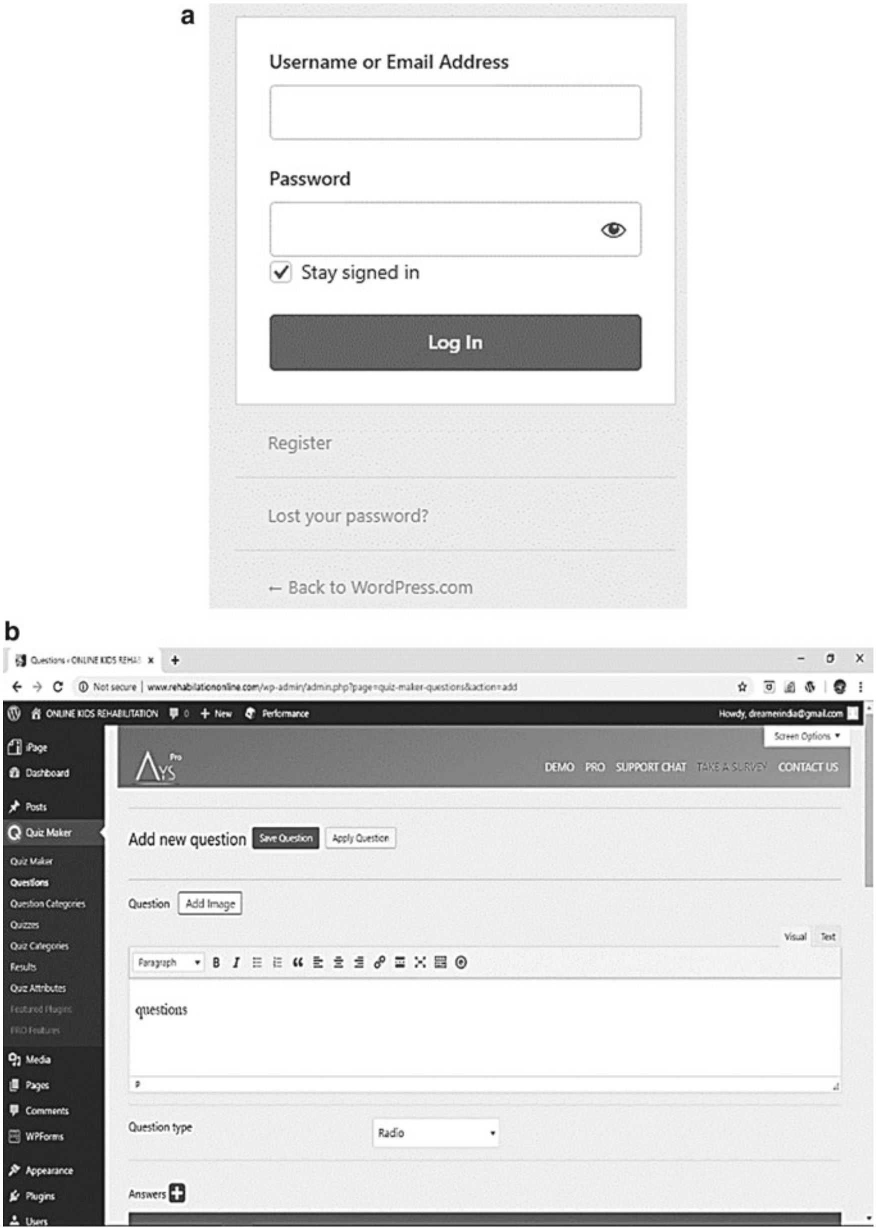
The questionnaire is framed by considering the age, computation time, math calculation, math fluency, applied problems, quantitative concepts and number series as parameters according to WJ IV. Some of the skills set questions used for identification of dyscalculia children are.

(a) Math Calculation—Basic arithmetic calculation such as addition, subtraction, multiplication and division.
(b) Math Fluency—Basic arithmetic calculation (Double and triple digits) within the stipulated timing.
(c) Applied problem—Analysis of questions to perform a suitable arithmetic operation.
(d) Quantitate concepts and Number series—Measures the knowledge about math symbols, series of numbers, identification of missing numbers.

Some of the sample math calculation, number comparison and place value questions for the screening of mathematical disability is shown in the Fig. 2a–c (Fig. 3).

4 Decision Tree Classification Methodology

Nearly 50 children's who are in the age of 8 and 9 years participated in the quiz. All the children were allowed to take the quiz in allotted timing. By using the proposed analytical tool, 50 datasets were generated. The datasets were formed by considering the parameters such as average time taken for each question and the aggregated score of each section, where the different sections are math calculation, math fluency, applied problem, Quantitate concepts and Number series. The block diagram of the proposed methodology is shown in Fig. 4.



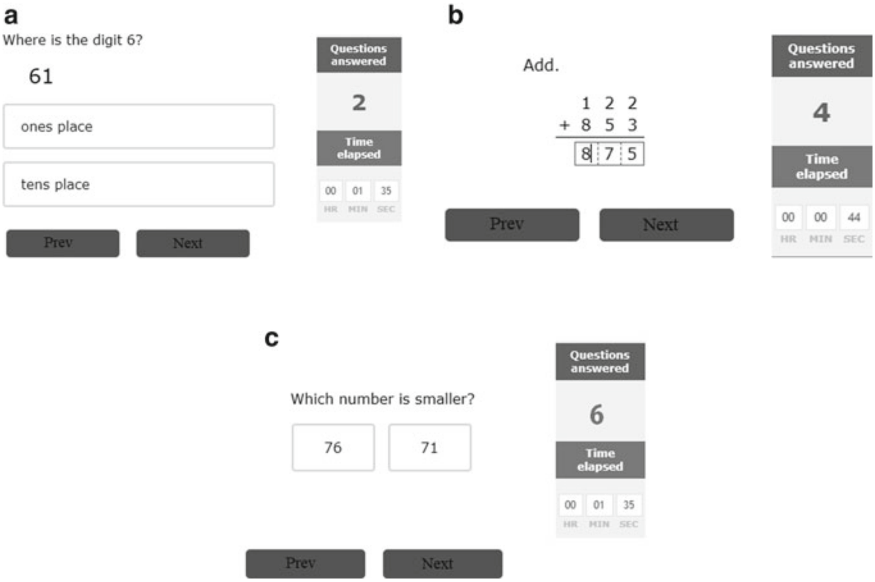
☑ Stay signed in

Not secure | www.rehabilitationonline.com/wp-admin/admin.php?page=quiz-maker-questions&action=add

Question: questions

Question type: Radio

Fig. 1 **a** Login Page **b** Page for adding a questionnaire



a

Where is the digit 6?

61

ones place

tens place

b

Add.

$$\begin{array}{r} 1\;2\;2 \\ +\;8\;5\;3 \\ \hline 8\;7\;5 \end{array}$$

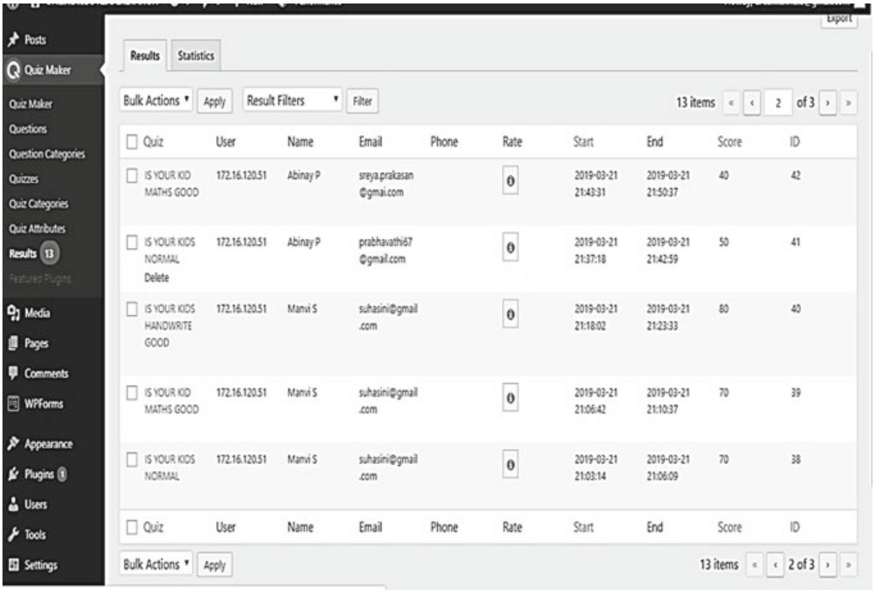
c

Which number is smaller?

76

71

Fig. 2 **a** Place value questions **b** Three-digit addition **c** Comparison question



☐ Quiz	User	Name	Email	Phone	Rate	Start	End	Score	ID
☐ IS YOUR KID MATHS GOOD	172.16.120.51	Abinay P	sreya.prakasan@gmai.com		0	2019-03-21 21:43:31	2019-03-21 21:50:37	40	42
☐ IS YOUR KIDS NORMAL Delete	172.16.120.51	Abinay P	prabhavathi67@gmail.com		0	2019-03-21 21:37:18	2019-03-21 21:42:59	50	41
☐ IS YOUR KIDS HANDWRITE GOOD	172.16.120.51	Manvi S	suhasini@gmail.com		0	2019-03-21 21:18:02	2019-03-21 21:23:33	80	40
☐ IS YOUR KID MATHS GOOD	172.16.120.51	Manvi S	suhasini@gmail.com		0	2019-03-21 21:06:42	2019-03-21 21:10:37	70	39
☐ IS YOUR KIDS NORMAL	172.16.120.51	Manvi S	suhasini@gmail.com		0	2019-03-21 21:03:14	2019-03-21 21:06:09	70	38
☐ Quiz	User	Name	Email	Phone	Rate	Start	End	Score	ID

Fig. 3 Test results in word press

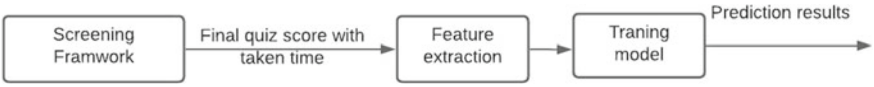


Fig. 4 Block diagram of proposed methodology

In the proposed model, the identification of children with dyscalculia is performed by decision tree classification. The decision is one of the simple and most familiar machine learning algorithms. The decision tree works accurately when the dataset is lesser in size. To test only important parameters (parameters which makes the most difference to classification) first, the decision tree is adopted. The data is filtered and sorted as the datasets are processed, and then datasets are used to identify the children with dyscalculia. The first step is to split the dataset into training and testing, the next step is to perform feature scaling. Then the model is fit in the decision tree. To predict a class label for an object, the decision tree algorithm processes datasets from the root of the tree. The root attribute values are compared with record attributes. Based on the similarity, the branch corresponds to the value and moved to the next node. The random forest [12] is introduced to eliminate the issue of overfitting problem in the decision tree. This method of combining the output of multiple individual models is called Ensemble Learning, where the random forest is an example of ensemble learning. The random forest allows combining several machine learning algorithms to provide better predictive results.

5 Results

The training dataset contains the data set on which the model can be trained, while the testing dataset involves data on which the testing model is applied to assess its accuracy. In the proposed framework, the data to be used as input is the questions set based on the Woodcock Johnson IV test. After training the model, the dyscalculia prediction accuracy is 95.3%. The accuracy plays a significant role as the result produced depends completely on this method. This makes the proposed tool to be reliable and accurately generates results without any manual work (Fig. 5).

6 Conclusion and Future Enhancement

The proposed knowledge-based tool efficiently and accurately generates the dataset to analyse and predict the children with dyscalculia problem. The proposed method is focussed to analyse the generated datasets effectively using decision tree machine learning algorithm. The proposed analytical tool is very interactive and identifies the children with dyscalculia in an efficient manner. The proposed tool will be useful for school teachers to avoid complex manual work carried out, and to identify the

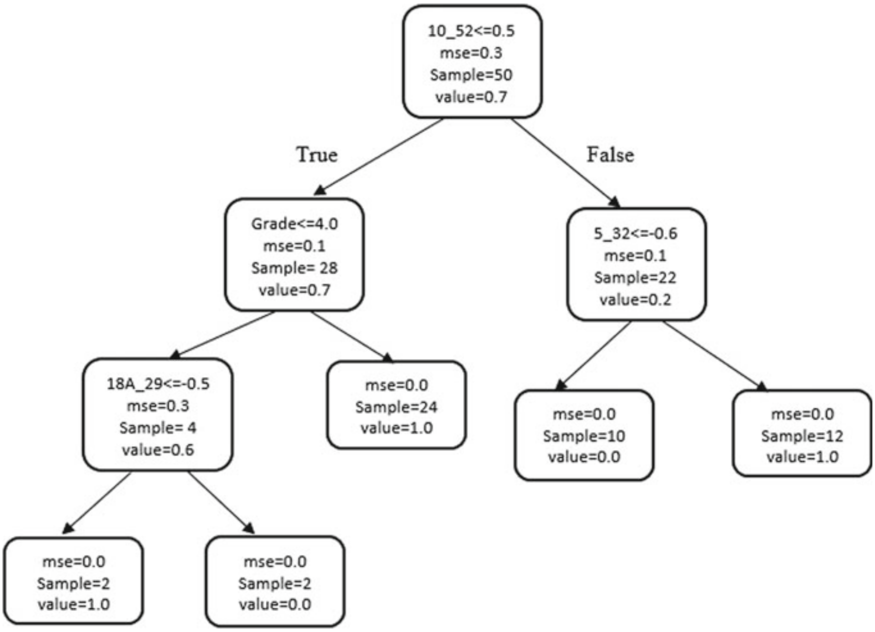


Fig. 5 Structure of Decision Tree based on the prediction result

children with dyscalculia at an early stage. In future, many parameters also can be included to identify the children with dyscalculia problem. The proposed tool can also be used to identify dyslexia and dysgraphia problems by considering the corresponding evaluation parameters (Table 1).

Table 1 Comparative Study based on different Tools

Name of the Author	Study conducted/ Tool	Results	Drawback
Mohd Syah [6]	Computerized play tool for dyscalculia intervention	Total post-test scores found that 57.9 percent of children in the intervention group did slightly higher than the control group during the five-day intervention period	Play model does not support local languages
Rani [7]	Web-based tool	Identify the level of learning capability of mathematics and allow the children to use appropriate the application	Does not suits for all mobile applications
Annemie [8]	Investigated some of the outcomes of studies for predicting the dyscalculia problem at the young children based on the behavioral signs	Dyscalculia has a strong heritability and that the likelihood of dyscalculia is enhanced by prematurity and very low birth weight. Additionally, in kindergarten, dyscalculia is frequently but not always accompanied by language deficits and impaired (procedural and conceptual) knowledge of counting, seriation, grouping, and quantity estimation	Questionnaire-based study
Dazhi Cheng at el [9]	İnvestigated whether deficits in cognitive visual perception were common to both dyslexia and dyscalculia	The findings revealed that both of dyslexia, dyscalculia and comorbid dyslexia with dyscalculia is marked by deficiencies in the understanding of numerosity and visual vision	Difficult to predict

(continued)

Table 1 (continued)

Name of the Author	Study conducted/ Tool	Results	Drawback
Rikard Ostergen [10]	Conducted three studies for examining the connection between cognitive abilities and arithmetic	The results suggested that the number of sensory deficits with memory functions together constitute risks for children with a mathematical learning disability	Suggesting some other basic development model based on von Asters and Shalev's model

References

1. David JM, Balakrishnan K (2009) Prediction of frequent signs of learning disabilities in school age Children using association rules, Proc Int Conf Adv Comput ICAC 2009, MacMillion Publishers Ltd., New York City, pp 202–207
2. https://www.epi.org/publication/bp278/
3. Handbook of Psychological Assessment [4 ed.] 0128022035, 9780128022030.
4. Gupta S, Kaur M, Lakra S, Dixit Y (2020) A comparative theoretical and empirical analysis of machine learning algorithms. Webology 17(1):377–397
5. Julie M, Balakrishnan K (2010) Significance of Classification techniques in prediction of learning disabilities. Int J Artif Intell Appl 1(4):111–120
6. Mohd Syah NE, Hamzaid NA, Murphy BP, Lim E (2016) Development of computer play pedagogy intervention for children with low conceptual understanding in basic mathematics operation using the dyscalculia feature approach. Interactive Learning Environ 24(7):1477–1496
7. Rani MFCA, Rohizan R, Rahman NAA (2014) Web-based learning tool for primary school student with dyscalculia. In: Proceedings of the 6th international conference on information technology and multimedia (pp 157–162) IEEE
8. Desoete A, Baten E (2017) Indicators for a specific learning disorder in mathematics or dyscalculia in toddlers and in kindergarten children. Belgian J Paedicatrics 19(2):122–124
9. Cheng D, Xiao Q, Chen Q, Cui J, Zhou X (2018) Dyslexia and dyscalculia are characterized by common visual perception deficits. Dev Neuropsychol. 43(6):497–507. https://doi.org/10.1080/87565641.2018.1481068 PMID: 29975105
10. Ostergen R (2013) Mathematics learning difficulties, cognitive conditions, development and predictions. Department of behavioral sciences and learning, ISSN 0282-9800
11. Devi A, Kavya G (2019) Intelligent system for identifying Dyscalculia based on Raspberry pi. Int Conf Commun Electron Syst (ICCES)
12. Random Forest Analysis in ML and when to use it - NewGenApps, NewGenApps (2020)

RSA Algorithm Using Performance Analysis of Steganography Techniques in Network Security

B. Barani Sundaram, N. Kannaiya Raja, Nune Sreenivas, Manish Kumar Mishra, Balachandra Pattanaik, and P. Karthika

Abstract The stable network has now become a prerequisite for every relationship. The security hazards are expanding and rendering wired/remote organizations and Internet providers fast, shaky and problematic. Today, safety efforts are all the more important in meeting the cutting edge demands of current emerging businesses. The need is also discussed in areas such as guards, where the key points of concern are safe and checked admission of properties identified with data security. In this paper, the author described the important measures and limits for the establishment of a secure organization with regard to enormous industry / authoritative needs. In providing remote organization access to different assets and interfacing various gadgets remotely, Wi-Fi networks are extremely basic. To deal with Wi-Fi hazards and organizational hacking attempts, different prerequisites are required. This article

---

B. B. Sundaram (✉)
ICT-CE, Computer Science Department, College of Informatics, Bule Hora University, Bule Hora, Ethiopia
e-mail: bsundar2@gmail.com

N. Kannaiya Raja
College of Informatics, Bule Hora University, Bule hora, Ethiopia
e-mail: kannaiyaraju123@gmail.com

N. Sreenivas
School of Electrical and Computer Engineering, Addis Ababa Institute of Technology, Addis Ababa University, Addis Ababa, Ethiopia
e-mail: ns_maruthi@yahoo.com

M. K. Mishra
Department of Computer Science, University of Gondar, Gondar, Ethiopia
e-mail: mishrasoft1@gmail.com

B. Pattanaik
Department of Electrical and Computer Engineering, Blue Hora University, Blue Hora, Ethiopia
e-mail: balapk1971@gmail.com

P. Karthika
Department of Computer Applications, Kalasalingam Academy of Research and Education, Krishnankoil, Viruthunagar, India
e-mail: karthikasivamr@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_56

discusses important protection efforts identified with different organizational conditions, so that an association may be set up with a fully assured organizational environment. Besides, a contextual investigation was examined to delineate the insignificant arrangement of measures necessary in any association to construct network security.

Keywords Steganography techniques · WAN · Advanced encryption standard (AES) · Open System Interface (OSI)

1 Introduction

Security of organizations can be defined as insurance of organizations and their administrations from unwanted change, annihilation, or disclosure, and assurance arrangements that the organization acts in basic circumstances and has no adverse effects on either customer or representative [1]. Besides, it involves arrangements made in a simple PC network foundation, techniques adopted by the director of the company to protect open assets for the organization and the organization from unauthorized clients. Every organization has data security plans these days, steganography assumes a significant part in ensuring the security of data of innovation applications [2]. Data security is a significant issue, for certain applications. The main concern, for example, web-based business banking, email, clinical data sets, thus some more, every one of them requires the trading of private data. For instance, let us consider an individual named Alice a sender who needs to send an information message which has a length of characters to a beneficiary called Bob. Alice utilizes an unstable correspondence channel. This could be a phone line, PC organization, or some other channel [3]. On the off chance that the message contains mystery information, they could be blocked and perused by programmers.

2 Literature Review

They characterized the centre functional systems' administration parts of security including PC interruption identification, traffic examination, and organization observing parts of organization security. A new methodology has been introduced for the controlled, shared use of dispersed security arrangements, called the security network, in which the gadget network ensures that a gadget is reliable and that interchanges between gadgets can be performed levelled out of the framework strategies [1]. Characterized data security is divided into three parts: information security, the security of the network framework and business security of the organization, and the business security model of the organization. In addition, a hypothetical rationale for safety protection for the programmed development system for large businesses has been created. A public key infrastructure (PKI)-based remote organization protection framework has been characterized [14]. Various instruments and treatments identified with cryptography and organizational security have been characterized in this [4–6]. Besides, the most recent problems found with network security innovation and its relevant implementations such as Advanced Encryption Standard (AES), CMAC validation mode and CCM mode for checked encryption guidelines are addressed in an extremely elaborate way [7]. These days, transferring data around an enterprise more safely and securely has become an important test for the company. The attacks and the organizational security activities characterize how a superior, sound and stable organization can be designed and maintained for an association/industry using organizational security devices [8]. This exploration focuses on the problems through which network security in an association can be supervised and maintained more proficiently. In addition to the security strategy and contextual investigation, the better management of the security control agency in an association would help a great deal to understand.

3 Performance Analysis of Steganography Using RSA Algorithm

The main advancement for a wide range of uses is platform and network technology. In current organizations, it is a fundamental prerequisite, and there is a noteworthy absence of security strategies that can be easily implemented. A "correspondence hole" exists between security innovation engineers and organizational designers. Organization configuration is a loop generated based on the model of the Open System Interface (OSI). When planning network defence, the OSI model has few priorities. It provides seclusion, ease of use, adaptability and convention normalization [9, 10]. To build stacks that allow unique turns of events, the conventions of different layers can be effectively joined. Instead of ensuring, the organizational configuration is not calculated (Fig. 1).

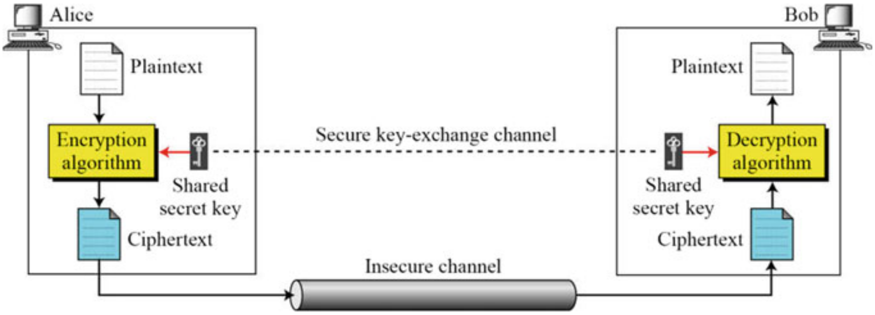


Fig. 1 Steganography technique model

There is no framework to fix the unpredictability of security specifications. When considering organizational protection, it should be stressed that the overall organization is secure. It is not just about the security of the PCs at each end of the correspondence chain. When travelling from one hub to the next hub, the correspondence channel should not be defenceless against attack. A programmer will focus on the channel of correspondence, obtain and decode the content, and re-integrate a print message, although making sure about the organization is similarly as significant as making sure about the PCs and scrambling the message.

On the off chance that P is the plaintext, C is the ciphertext, and K is the key,

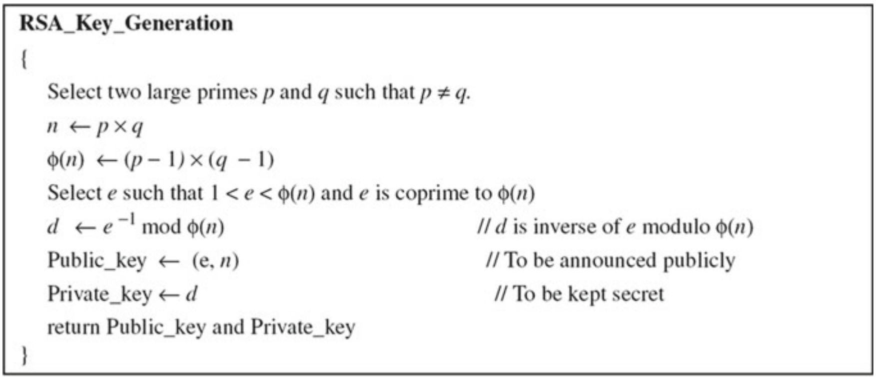
$$\text{Encryption} : \mathrm{C} = \mathrm{E_k(P)} \quad \text{Decryption} : \mathrm{P} = \mathrm{D_k(C)} \tag{1}$$

$$\text{In Which,}\ \mathrm{D_k(E_k(x))} = \mathrm{E\,_k(D\ k\ (x))} = \mathrm{x} \tag{2}$$

3.1 RSA Algorithm

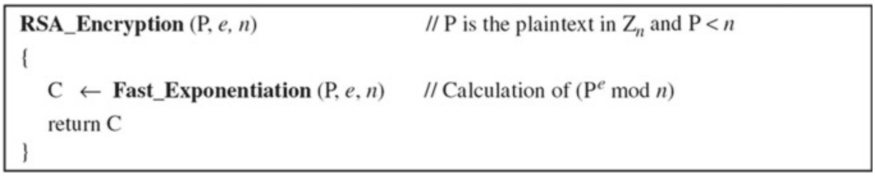
RSA represents Rivest Shamir and Adleman name of three innovators. RSA is one of the primary functional public key cryptosystems and is generally utilized for secure information transmission. In such a cryptosystem, the encryption key is public and varies from the decoding key which is left well enough alone. In RSA, this deviation depends on the functional trouble of calculating the result of two huge prime numbers, the considering issue. RSA represents Ron Rivest, Adi Shamir and Leonard Adleman, who first openly depicted the calculation in 1977. RSA includes a public key and a private key. The public key can be known by everybody and is utilized for scrambling messages. Messages scrambled with the public key must be unscrambled in a sensible measure of time utilizing the private key. The keys for the RSA calculation are produced the accompanying way.

Algorithm 1



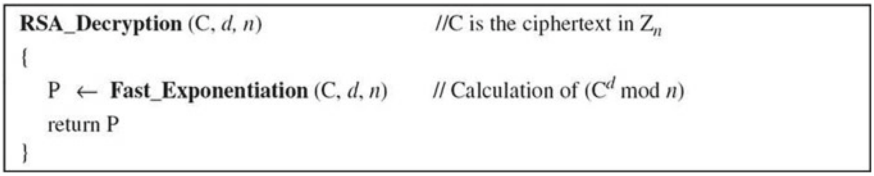
```
RSA_Key_Generation
{
    Select two large primes $p$ and $q$ such that $p \neq q$.
    $n \leftarrow p \times q$
    $\phi(n) \leftarrow (p - 1) \times (q - 1)$
    Select $e$ such that $1 < e < \phi(n)$ and $e$ is coprime to $\phi(n)$
    $d \leftarrow e^{-1} \bmod \phi(n)$                // $d$ is inverse of $e$ modulo $\phi(n)$
    Public_key $\leftarrow$ (e, $n$)                  // To be announced publicly
    Private_key $\leftarrow d$                         // To be kept secret
    return Public_key and Private_key
}
```

Algorithm 2



```
RSA_Encryption (P, e, n)                    // P is the plaintext in $Z_n$ and $P < n$
{
    C ← Fast_Exponentiation (P, e, n)       // Calculation of ($P^e$ mod n)
    return C
}
```

Algorithm 3



```
RSA_Decryption (C, d, n)                    //C is the ciphertext in $Z_n$
{
    P ← Fast_Exponentiation (C, d, n)       // Calculation of ($C^d$ mod n)
    return P
}
```

The world is getting more interconnected of the Internet and new system administration innovation. There is enormous measure of individual, military, business and government data on system administration foundations overall accessible. Organization security is happening to extraordinary significance on account of licenced innovation that can be handily procured through the Web.

4 Result Analysis of Network Security

Organization security begins with approval, normally with a username and a secret key. Organization security comprises the arrangements and approaches received by an organization manager to forestall and screen unapproved access, change in framework, abuse or disavowal of a PC organization and organization available assets (Fig. 2).

Fundamentally, network security includes the approval of admittance to information in an organization, which is constrained by the organization administrator. It has gotten more critical to PC clients and associations. If this is handled, then the firewall security arrangements will come into effect, for example, what administrations are permitted to be gotten to for network clients (Fig. 3).

So that to forestall unapproved admittance to a framework, this segment may neglect to check possibly hurtful substance; for example, PC worms or Trojans being communicated over the organization against infection programming or interruption identification frameworks (IDS) help recognize the malware. Today, irregularity may likewise screen the organization like Wireshark traffic and might be logged for review purposes and for later on elevated-level examination in the framework. Correspondence between two hosts utilizing an organization might be utilized encryption to keep up security strategy.

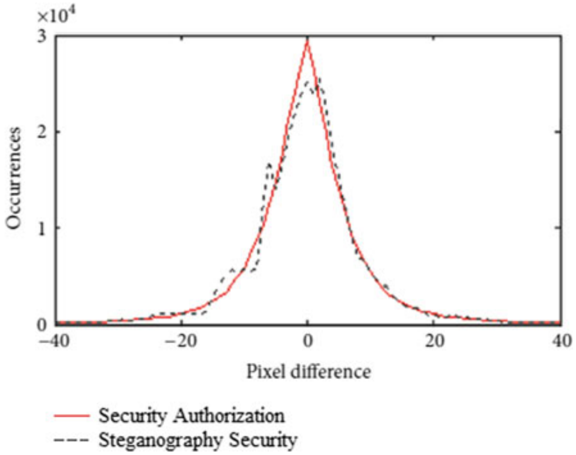


Fig. 2 Network security includes the approval of access

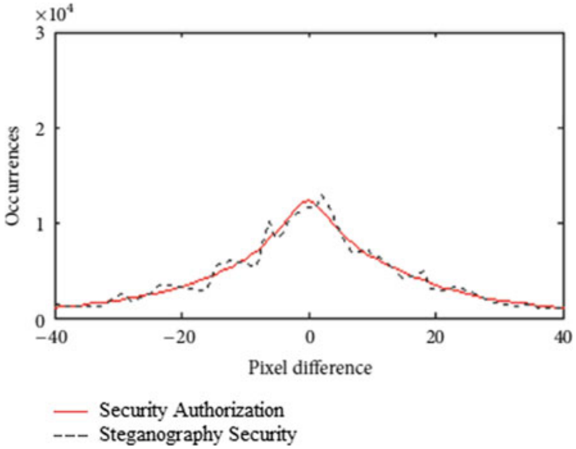


Fig. 3 To check possibly forestall approved admittance to a framework

5 Conclusion

Security is a significant issue besides huge associations. There are different concepts and proposals for protection and risk controls. The safety efforts ought to be planned and given; initially, an organization should know its need for security on various degrees of the association, and later, it ought to be actualized for various levels. Security arrangements ought to be planned first before its execution in such a manner, with the goal that adjustment of selection can worthy with effectively sensible. The security framework should be secure, and also, the end-user should be adaptable to satisfy everyone. The user must believe that the security framework should be appropriate.

References

1. Karthika P, Vidhya Saraswathi P (2017) Content based video copy detection using frame based fusion technique. J Adv Res Dyn Control Syst 9:885–894
2. Karthika P, Vidhya Saraswathi P (2020) Raspberry Pi—A tool for strategic machine learning security allocation in IoT, Apple Academic Press/CRC Press (A Taylor & Francis Group). Proposal has been accepted (provisionally) for the book entitled "Making Machine Intelligent by Artificial Learning", to be published by CRC Press
3. Depren O, Topallar M, Anarim E, Ciliz MK (2019) Anomaly and Misuse detection in machine learning schemes for sensor networks. Expert Syst Appl 29(4):713–722
4. Karthika P, Vidhya Saraswathi P (2020) IoT using machine learning security enhancement in video steganography allocation for Raspberry Pi. J Ambient Intell Human Comput https://doi.org/10.1007/s12652-020-02126-4 Impact Factor 1.91
5. Anuar NB, Shamshirband S, Rohani VA, Kiah MLM, Petkovic D, Misra S (2019) Artificial immune system for detecting intrusion in stego-crptonetworks. J Comput Netw Appl 42(4):102–117
6. Maleh Y, Ezzati A, Qasmaoui Y, Mbida M (2020) Intrusion detection system fora global geometric image using machine learning system. In:Fifth International Symposium on Frontiers in Ambient and Mobile Systems (FAMS 2015) in association with Elsevier-Procedia Computer Science, vol 52, pp 1047–1052
7. Karthika P, Vidhya Saraswathi P (2017) A survey of Content based Video Copy detection using Big Data. Int J Sci Res Sci Technol (IJSRST). 3(5): 114–118. Online ISSN : 2395–602X, Print ISSN : 2395–6011, May-June 2017. https://ijsrst.com/ICASCT2519
8. Ganesh Babu R, Elangovan K, Maurya S, Karthika P (2020) Multimedia security and privacy on real time behavioral monitoring in Machine Learning IoT application using Big Data Analytics. In: Raghvendra K, Sharma R, Pattnaik PK (eds) Multimedia technologies in the Internet of Things Environment (pp 155–177) Springer
9. Shen Y, Liu S, Zhang Z (2020) Detection of hello flood attack caused by malicious cluster heads on efficient machine learning technique. Int J Adv Comput Technol 7(2):40–47
10. Karthika P, Vidhya Saraswathi P (2019) Digital Video Copy detection using steganography frame based fusion techniques. In: International Conference on ISMAC in Computational Vision and bio-engineering. pp 61–68. https://doi.org/10.1007/978-3-030-00665-5_7

Performance of Two-Link Robotic Manipulator Estimated Through the Implementation of Self-Tuned Fuzzy PID Controller

Aditi Saxena, Rishabh Chaturvedi, and Jitendra Kumar

Abstract Robot manipulators become intensified major segments in the production sectors that are utilized in numerous applications such as welding, granulating, mechanical dealing, and assembling because of velocity, accuracy, and repeatability. These applications necessitate a legitimate path plan, appropriate generation of direction, and above all a control design. In a manipulator system, the links are expanded based on efficiency and versatility is upgraded alongside the complexity. This controller tuned to produce the least weighted sum of integral of absolute error and component of positive change in controller output. The controller has the advantage of examining and controlling a highly non-linear and rigid manipulator system. To optimize the system, a genetic algorithm is being used to provide an actual and desired result by incorporating a combination of a genetic algorithm with self-tuning fuzzy. PID is composing the controller much more superior and reliable.

Keywords Self-tuned fuzzy PID · Manipulator · Genetic algorithm

1 Introduction

The term robot is basically a Czech word that means slaved laborers, and it came into existence in 1920. Robots acts like a machine and is dominated by a controller in a robot. Further, the structure of the robot is replicated as similar to human body parts like a head, two arms, and two legs, waist, and so on. Industry-specific robots play out a few tasks, for example, picking and fixing objects, and movement adjusted from seeing how comparable manual tasks are taken care of by a completely working human arm. Such automated arms are otherwise called robotic manipulators. The proposed work aims to have innovation is a multifaceted concept that includes modifications to products, alterations to organizational aspects dependent on innovations, strategies, rebuilding of organizational culture, and revisions to production processes. Accordingly, such optimized controlled robotic manipulator is

---

A. Saxena (✉) · R. Chaturvedi · J. Kumar
IET Department of Electronics and Communication Engineering, GLA University, Mathura 281406, India
e-mail: aditeesaxenaa29@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_57

appropriate for process innovation since they authorize organizations to concentrate on unique processes for the production of products and services. The second factor is the organizations today are confronted with expanding work costs and a deficiency of laborers, and are thus spending in robotics. Robots nevermore demand promotions and can work nonstop. The third factor is the pressure to increase the production rates to compete the market, and the fourth factor is the increased productivity. The fifth factor explains about the repeatability where the robot's drive product quality or consistency and reduces waste. The last factor represents the speed in which the robotics assist in the increased production and hence reduction in the waiting time.

This controller can easily accomplish all these factors, where the controller is used for underwater controlling, etc. The proposed model can move and grip devices with guidance which can be operated through various programmed motions as it is both programmable and multifunctional devices [1]. Robotics is a combination of two branches [2, 3] that are electronics through the field of control, manufacturing, designing, and kinematics which help in positioning and orientation of the manipulator devices having multi-degree liberty can be positioned with the help of these manipulators. High and improved control strategies are required to achieve accuracy in the trajectory tracking of the manipulator, dynamic model of a robotic manipulator also play a very important role; it is responsible for the performance of a robotic manipulator. To achieve the desire performance and response of the system, a new design is needed for an accurate mathematical model, and after this, the various control strategies are applied to get precise trajectory tracking [4]. For industry purpose, the manipulators that are being used are highly non-linear and coupled; a second-order non-linear differential equation is being used to design a dynamic model for two links rigid robotic manipulator [5]. If the number of links increased simultaneously complexity also increased in terms of position and velocity of the manipulator. To overcome this challenging task, a control system is employed to monitor the manipulator [6]. Especially after designing a dynamic model, it is difficult to attach a precise and powerful controller to the plant that could be able to handle the non-linearities and many other unknown parameters. The controller is managed by adopting a tuning algorithm to have flexibility over the automation, where the entire world is searching for automatic processing in every aspect of life. Many different types of the algorithm are available in current trends such as particle swarm optimization (PSO), optimization method of ant colony [7], cuckoo search algorithm (CSA) [8, 9], genetic algorithm (GA), and so on. In the proposed system, genetic algorithm is used to produce high-quality solutions for optimization and search issues by depending on the process of natural selection such as crossover, mutation, crossover, and selection. The main aim is to reduce the amount of error. The controller will operate with the self-tuned fuzzy logic controller to have the least value of error and can also supervise the non-linearities [10–12]. The limitations of the proposed work are that this controller could be able to control a two-link manipulator having a payload of 0.5 kg [13] (Fig. 1 and Table 1).

2 Dynamic Model for Manipulator

In this study, the configurations of the robotic manipulator is explained and dynamic model was designed by studying the equations of motions for manipulator arm which was given by Lin [14],

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}\begin{bmatrix} \theta_{11} \\ \theta_{22} \end{bmatrix} + \begin{bmatrix} P_{11} \\ P_{21} \end{bmatrix} + \begin{bmatrix} f_{r1} \\ f_{r2} \end{bmatrix} + \begin{bmatrix} fn_{1p} \\ fn_{2p} \end{bmatrix} = \begin{bmatrix} \tau f_{1p} \\ \tau f_{2p} \end{bmatrix}$$

where

$$\begin{aligned} S_{11} = {} & I_{1p} + I_{1p} + m_{11} l_{c1}^2 + m_{22} l_{c2}^2 + 2 m_{22} l_{11} l_{c2} \cos\theta_{22} \\ & + m_{vp} l_{11}^2 + m_{vp} l_{22}^2 \\ & + 2 m_{vp} l_{11} l_{22} \cos\theta_{22} \end{aligned} \tag{1}$$

$$S_{12} = I_{2P} + m_{22} l_{c2}^2 + m_{22} l_{11} l_{c2} \cos\theta_{22} + m_{vp} l_{22}^2 + m_{vp} l_{11} l_{22} \cos\theta_{22} \tag{2}$$

$$S_{21} = S_{12} \tag{3}$$

$$S_{22} = I_{2P} + m_{22} l_{c2}^2 + m_{vp} l_{22}^2 \tag{4}$$

$$P_{11} = -m_{22} l_{11} l_{c2} (2\theta_{11} + \theta_{22}) \theta_{22} \mathrm{Sin}\theta_{22} - m_{vp} l_{11} l_{22} (2\theta_{11} + \theta_{22}) \theta_{22} \sin\theta_{22} \tag{5}$$

$$P_{21} = m_{22} l_{11}\, \theta_{11}^2 l_{c2} \sin\theta_{22} + m_{vp} l_{11}\, \underset{11}{\theta^2}\, l_2 \sin\theta_{22} \tag{6}$$

$$f_{r1 = } b_{1vp}\, \theta_{11} \tag{7}$$

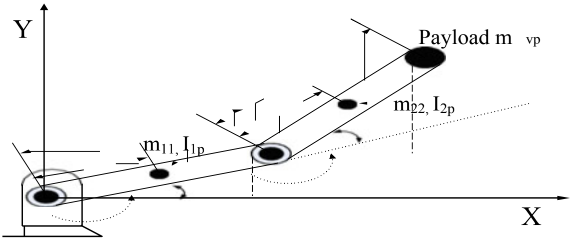
$$f_{r1 = } b_{1vp}\, \theta_{22} \tag{8}$$

$$\begin{aligned} fn_{1p} = {} & m_{11} l_{c1} g \cos\theta_{11} + m_{22} g (l_{c2} \cos(\theta_{11} + \theta_{22}) \\ & + l_{11} \cos\theta_{11} + m_{vp} g (l_{22} \cos(\theta_{11} + \theta_{22}) \\ & + l_{11} \cos\theta_{11}\ (9) \end{aligned} \tag{9}$$

$$fn_{2p} = m_{22} l_{c2} g \cos(\theta_{11} + \theta_{22}) + m_{vp} l_2 g \cos\ (\theta_{11} + \theta_{22}) \tag{10}$$

$$\theta_{11} = \frac{\tau f_{1p} - f_{n1p} - f_{r1} - P_{11} - S_{12} * \theta_{22}}{S_{11}} \tag{11}$$

Fig. 1 Diagram of two link robotic manipulator



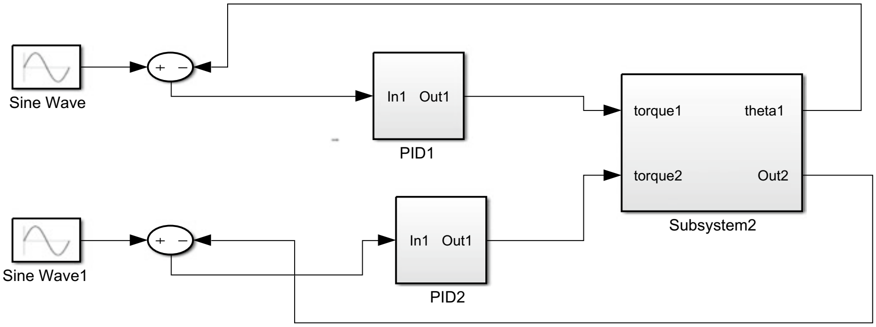


Fig. 2 PID structure

$$\theta_{22} = \frac{\left(\tau f_{2p} - f_{n2p} - f_{r2} - P_{21} - S_{12} * \theta_{11}\right)}{S_{22}} \tag{12}$$

where θ_{11} and θ_{22} shows the links position; τf_{1p} and τf_{2p} are the generated torque; f_{r1} and f_{r2} are the dynamic friction coefficients; m_{11} and m_{22} shows the masses. l_{11} and l_{22} shows the length, I_{1p}, I_{2P} shows inactivity.

Initially, the mathematical model for each equation is constructed as S_{11},S_{12},S_{21},S_{22},P_{11},P_{21},f_{r1},f_{r2},fn_{1p},fn_{2p}. Further, the mathematical model is again converted into a dynamic model for further processing. On the MATLAB SIMULINK, a subsystem for these equations is created and appended with the equations shown in no 11 and 12. This whole mathematical model accommodates various controllers named subsystem 2 that are given in Fig. 2 (Figs. 3, 4, 5 and 6).

2.1 Various Controller Structure

2.1.1 PID Controller Structure

This controller structure is designed using the MATLAB SIMULINK. The design of the simple PID controller is constructed by connecting link 1 and link 2. These

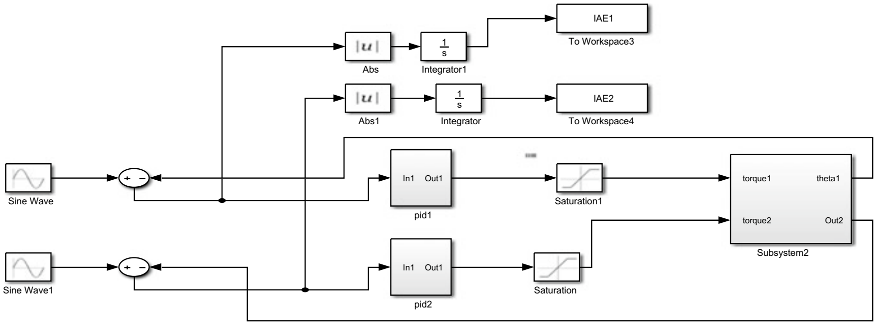


Fig. 3 Optimized PID structure

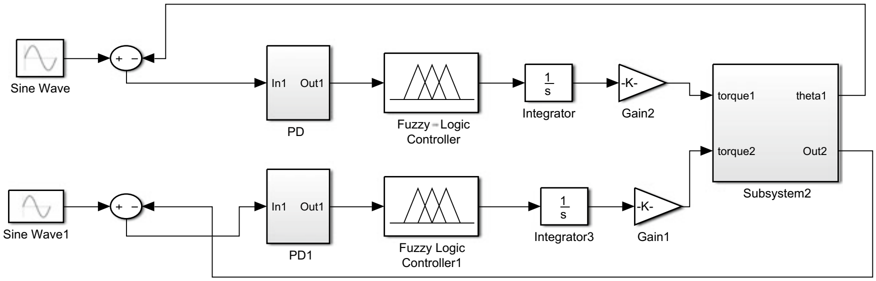


Fig. 4 Fuzzy PD structure

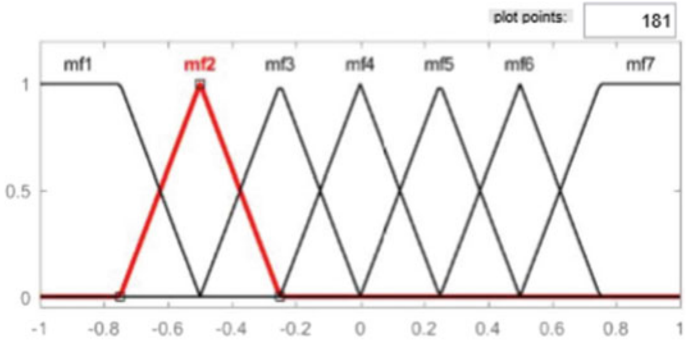


Fig. 5 Membership function plot

links reflect the operations of the equations defined in the mathematical model. Subsequently, a subsystem is being created of the model. The control boundaries of the PID structure are namely, proportional term, integral term, and derivative term. The proportional term refers to a general control activity proportional to the error signal through the steady gain factor. The integral term aims to decrease consistent

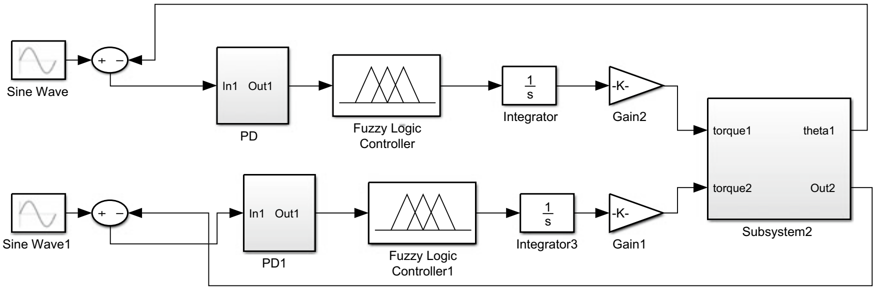


Fig. 6 Fuzzy PI structure

state errors through low-frequency remuneration by an integrator. The derivative term aid to improve transient reaction through high-frequency remuneration by a differentiator. Each link has its own PID controller connected for generating the optimal output. The inputs supplied for each link are in terms of the sine wave, and the controller supports in all sorts of processing, and the final output is matched with the mathematical model equations. If any mismatch occurs between the corresponding desired value and the actual value, then the controller is again fine-tuned towards the gain factor. The genetic algorithm is employed to enhance time and achieve higher precision accuracy.

3 Genetic Algorithm for Optimization

Nature has consistently been an incredible wellspring of motivation for all humankind. The genetic algorithm is associated with search calculations based on the ideas of selection and genetics. In GAs, a pool or a populace of potential clarifications are given for the specified issue. These clarifications at that point go through recombination and transformation (like in genetics), delivering new solutions, and the cycle is rehashed over different circumstances. Every value is specified as a fitness value, and the fitter is given a higher opportunity to yield more fitter. An optimistic approach such as integral absolute error (IAE) is determined to eradicate the number of errors. To operate with the IAE functionality, a small piece of code is written in the command window. The next process is to assign a sim function inside the MATLAB, which is being inbuilt under the optimization tool. By executing the command as optimtool in the command window a pop-up window gets activated, and inside the textbox type the term solver. Finally, run the GA commands with the number of variables associated with the gains. The sequence of steps involved in the genetic algorithm is (1) Population, (2) Fitness scaling, (3) Selection, (4) Reproduction, (5) Mutation, (6) Crossover, and (7) Migration. The population comprises of all the data. The scaling function converts raw fitness scored returned by the fitness function in a given range. The selection function determines parents for next-generation based on their scaled values. The reproduction essentially demonstrates the functionality of creating children for each new generation then progressed by the mutation function, which makes a small random change in the individual which provides genetic diversity. The crossover consolidates two individuals or parents to form a new individual for the next-generation. The values for each operation are assigned using the two commands such as assignin and sim assignin which are defined inside the workspace. The sim function returns a single simulink (simulation output object) that contains all the simulation output. The obtain results aids to plot the output signal value against the time.

3.1 PID Optimized Controller Structure

The genetic algorithm technique that is discussed in the previous section is being used in this model for optimization. This is the reason for the optimized PID controller design; IAE is being attached to the controller input. The rest of the whole structure is the same as the PID controller structure, where it contains only two blocks, and the initial block consists of absolute, and the other is integrator over time. The IAE integrates absolute error value and provides the least value of error.

3.2 Fuzzy PD Controller Optimized Structure

PID controller is alone insufficient to handle all the non-linearities, so there exists a need for adding the fuzzy logic functionalities. The fuzzy controller or logic consists of membership functions which are described in Fig. 7. The membership functions are being shown each has their respective ranges as fuzzy only accept 0 or 1; so accordingly, the range is being decided with the help of a rule base a fuzzy logic. The design is created by typing fuzzy in the command window, a fuzzy logic appears and it is possible to call this logic to the workspace where our structure has been

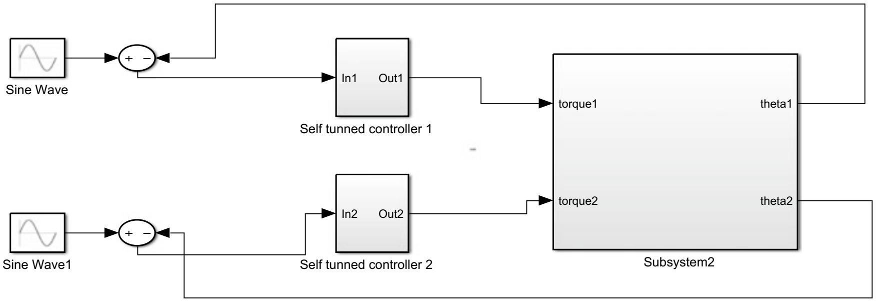


Fig. 7 Optimized self-tuned (FUZZY PID) model in SIMULINK

designed. Further, the fuzzy logic is being combined with PD controller and termed as fuzzy PD controller (Fig. 8).

3.3 Fuzzy PI Optimized Structured

This structure is the same as the above one the main difference here is the PD controller is being integrated which results in PI controller, and fuzzy logic is also being attached between PD and integrator. The whole structure as a sum give origin to fuzzy PI controller for optimization, again genetic algorithm is being used with the help of inbuilt optimtool function.

3.4 Optimized Self-Tuned Fuzzy PID (STFPID) structure

Here, two fuzzy layer are utilized, the output of fuzzy layer 1 is being multiplied by the product of gains and output of fuzzy layer 2, then fed to the plant. The fuzzy logic self-tuning structure is being represented below is termed as a self-tuned fuzzy PID controller (STFPID). This controller gives the lesser error value after optimizing the genetic algorithm.

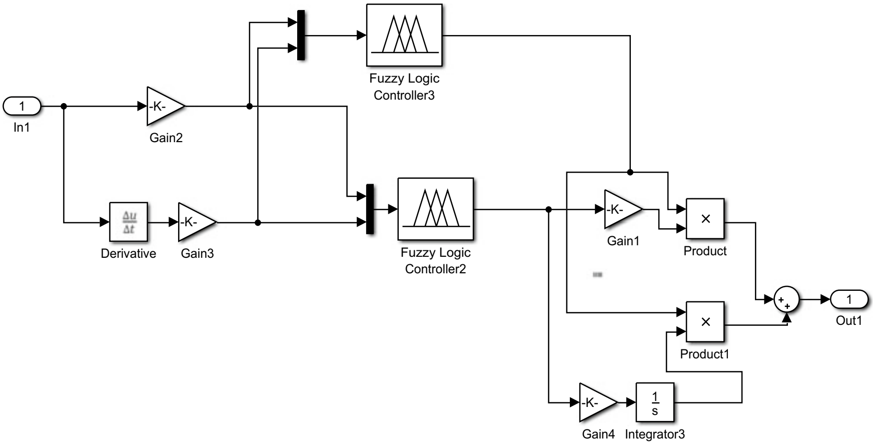


Fig. 8 Structure of self tuned fuzzy PID controller

4 Results and Discussions

For tracking the trajectory, the two-link robotic manipulator and the controller are optimized on its own with the help of a self-tuned fuzzy PID controller. The controller structure as discussed above, these controllers is working efficiently and fulfilling all their aims for which they have been designed. To check all this, a simulator is needed and MATLAB is used to compare each optimized controller. Figure 9 represents the tracking, the reference, and the output wave is perfectly matched. If it does not match, differentiate it between these two waves with the guidance of color-coding. The dynamic model is designed with the equations, and the controllers, which are being attached to it, are perfect and working precisely. Figure 10 describes a self-tuned fuzzy PID controller when operating through the genetic algorithm process as a simulation result, it shows the minimum error value. In comparison to all the other controllers which are being used, this controller contains many features. The self-tuned structure is allowing the controller to automatically tune itself depending upon the presence of non-linearities. The fuzzy has many advantages and can deal with multiple-links non-linearities, increased time, overshoot, and many more. The base

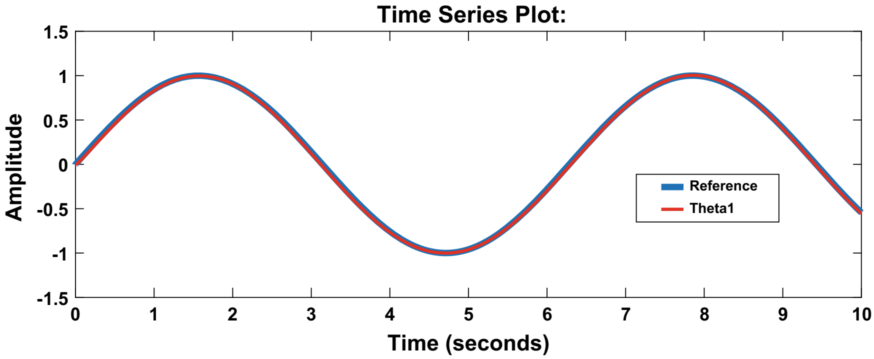


Fig. 9 Representing accurate tracking

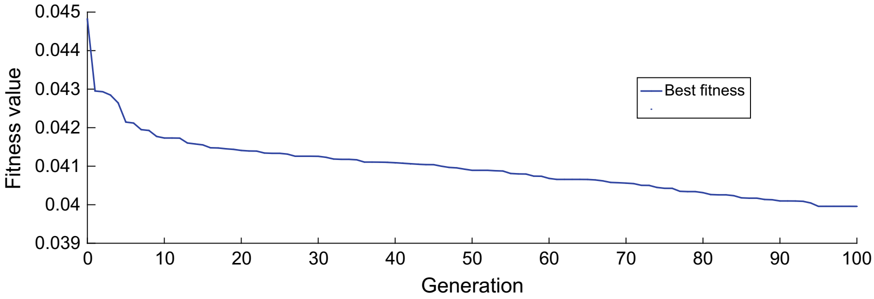


Fig. 10 Self-tuned fuzzy PID after optimization

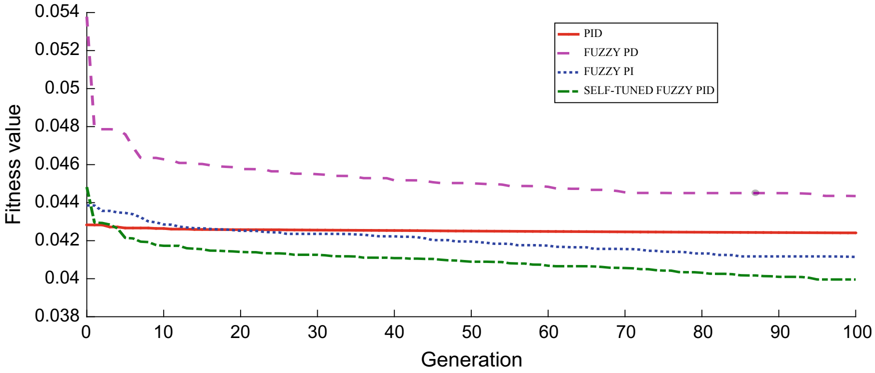


Fig. 11 IAE values of various controllers

for the fuzzy is PID in which all the three control parameters such as proportional, integral, and derivative are combined and serves in reducing error and providing stability to the system. Figure 11 showing the simulation result after working through a genetic algorithm all the five controllers are being individually simulated up to 100 generations. The self-tuned fuzzy PID controller has the least value of error as compared to any other. Table 2 represents the mathematical value of error for each controller in descending order and the self-tuned fuzzy logic has an error value of 0.03665 which is like next to perfection in the third graph all the controller's error. These are identified by separate graphs that are being differentiated with different color coding. This controller can work as an adaptable controller and can self-tuned and optimized itself according to the surrounding parameters and non-linearities.

5 Conclusion

In today's world of automation, the need for developing an adaptable controller is essential. This controller is working as a self-adaptable controller that can be tuned and adjusted on its own with the help of a self-tuned fuzzy PID controller and genetic algorithm. With the optimized simulation results, it is found that this controller has the least value of error when compared to other controllers that are also being designed and simulated. The proposed controller not only helps to control the two link manipulator but also has the application for controlling the underwater robot, to control an adaptable gripper, and many more.

Table 1 Parameters of the dynamic model

Parameters	Link 1	Link 2	Values
m_{11}	0.392924 kg	m_{22}	0.094403 kg
l_{c1}	0.104648 m	l_{c2}	0.081788 m
l_{11}	0.2032 m	l_{22}	0.1524 m
I_{1P}	0.0011411 kg m^2	I_{2P}	0.0020247 kg m^2
b_{1vp}	0.141231 N-m/radian/s	b_{2vp}	0.3530776 N-m/radian/s
m_{vp}	0.56699 kg		
g	/s^2		

Table 2 Tabulated form of IAE value in descending order

Various controller	Values of IAE
PID without optimization	0.2218
PID with optimization	0.0429
Fuzzy PD with optimization	0.04438
Fuzzy PI with optimization	0.04114
Self-tuned Fuzzy PID using GA	0.03665

References

1. Sharma R, Rana KPS, Kumar V (2014) Performance analysis of fractional order fuzzy PID controllers applied to a robotic Manipulator. Expert Syst Appl 41(9):725, 4274–4289:726
2. Kumar V, Rana KPS, Mishra JK, Nair SS (2016) A robust fractional order fuzzy P+ fuzzy I+ fuzzy D controller for nonlinear and uncertain system. Int J Autom Comput 14(4):474–488
3. Kumar V, Rana KPS (2017) Nonlinear adaptive fractional order fuzzy PID control of a 2-link planar rigid manipulator with payload. J Franklin Inst 354:993–1022
4. Kumar J, Kumar V, Rana KPS A fractional order Fuzzy PD+I controller for three link electrically driven rigid robotic manipulator system. J Intell Fuzzy Syst IOS Press, Netherlands (SCI Index, Impact factor-1.261)
5. Weile DS, Michielssen E (1997) Genetic Algorithm Optimization applied to electromagnetics, a review. IEEE Trans Antennas Propag 45(3):343–353
6. Kong Z, Jia W, Zhang G, Wang L (2015) Normal parameter reduction in soft set based on particle swarm optimization algorithm. Appl Math Model 39:4808–4820
7. Ghanbari A, Kazemi SMR, Mehmanpazir F, Nakhostin MM (2013) A cooperative ant colony optimization-genetic algorithm approach for construction of energy demand forecasting knowledge based expert systems. Knowl-Based Syst 39:194–206
8. Jagatheesan K, Anand B, Samanta S, Dey N, Ashour AS, Balas VE (2017) Design of a proportional-integral-derivative controller for an automatic generation control of multi-area power thermal systems using firefly algorithm, IEEE/CAA J Automatica Sinica, 1–14 (2017). https://doi.org/10.1109/JAS.2017.7510436
9. Yang XS, Deb (2009) Cuckoo search via Lévy Flights In: . Proceedings world congress on nature and biologically inspired computing, India, pp 210–214
10. Yang XS, Gandomi AH (2012) Bat-algorithm, a novel approach for global engineering optimization. Eng Comput 29(5):464–483
11. Ohtani Y, Yoshimura (1996) Fuzzy control of a manipulator using the concept of sliding mode. Int J Syst Sci 27(2):179–186

12. Hazzab A, Bousserhane IK, Zerbo M, Sicard P (2006) Real-time implementation of fuzzy gain scheduling of PI controller for induction motor machine control. Neural Process Lett 24:203–215
13. Sharma R, Bhasin S, Gaur P, Joshi D (2019) A switching-based collaborative fractional order fuzzy logic controllers for robotic manipulators. Appl Math Modell
14. Lin F (2007) Robust control design: an optimal control approach. John Wiley & Sons Ltd., England

Rule Based Part of Speech Tagger for Arabic Question Answering System

Samah Ali Al-azani and C. Namrata Mahender

Abstract Part of Speech (POS) Arabic wording is difficult to read in detail and its functionality affects many programs and activities in the Natural Language Processing (NLP) area. POS tagging is a process to assign POS such as a verb, adjective, adverb, noun in each word for any sentence. Farasa is an active and reliable text processing toolkit for Arabic documents. It is an assortment of Java libraries and CLIs for MSA.2. These incorporate a separate tool for Arabic text Diacritizer, segmentation/tokenization module, POS tagger, Named Entity Recognition (NER), and parsing. One of the limitations over the Farasa affects the post-processing results due to the presence of inappropriate tags. For our application on question answering system(QAS) correct POS, tagging is essential for better accuracy. The POS tagger is developed using a rule-based approach which is based on domain-specific. The corpus (database) on which the rule-based POS tagger is built is centered on the core subject of the Arabic language 4th standard textbook of Arabic Medium state board of Yemen. During the development of QAS, the POS tagger is a very essential stage in which the answers for the framed questions are obtained from the paragraphs of a given lesson. The present article provides insights into the complete process of linguistic rule-based POS tagger development for QAS. Sentence segmentation, word tokenization, to stemmer development which becomes an important part of proper morphological analysis is explained. As a result, the morphological analyzer is the input to the rule-based POS tagger. Ultimately, in this article, a comparison of marking based on our POS rule with Farasa is presented and for QAS, our rule-based POS tagger gave better results than Farasa.

Keywords Arabic language · Natural language processing (NLP) · Part of speech (POS) tagging · Morphological analyzer · Stemmer · Tokenization

---

S. A. Al-azani (✉) · C. Namrata Mahender
Department C.S. and I.T, Dr. Babasaheb Ambedkar, Marathawada University, Aurangabad, Maharashtra, India
e-mail: alazani183@gmail.com

C. Namrata Mahender
e-mail: nam.mah@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_58

1 Introduction

Natural language processing (NLP) is a category of artificial intelligence that assists computers to recognize, translate, and administer human language. It is utilized in numerous disciplines and incorporates computational linguistics, computer science, and so on. It also bridges the gap between computer understanding and human communication [1]. This methodology recognizes natural language texts and linguistic procedures namely, part of speech (POS) tagger, lexicon, and tokenization. These are applied to recreate inquiries into the right inquiry that separates the significant answers from a structured database [2]. The inquiries dealt with by this methodology are of Factoid type and have a profound semantic understanding [3].

1.1 Arabic Natural Language Processing

The Arabic language is a combination of many variations among different similarities that have a specific event, such as the standard official text of media and education throughout the Arabian World [4]. It is one of the top languages in the world. Its phenomenal script, distinguished style, and strong vocabulary confer a unique character and characteristic to the language. Arabic is the largest member of the Semitic language family. Nowadays, it is an official language in more than 20 countries and over 300 million native speakers [5]. Concerning other Semitic dialects, Arabic morphology was set up around the abstract idea of the root, three consonants articulating to significance, regardless of whether exact or ambiguous. The conventional derivational morphology is dependent on the root-and-pattern model concentrating on this abstract consonantal root. A pattern is an intermittent affix (or transfix) composed of vowels and non-revolutionary consonants embedded around spaces for the root consonants. To each pattern, conventional syntax relates a morphological classification or potentially inflectional highlights. These formalizations are applied by traditional grammar to depict both inflectional and derivational morphology. Moreover, the quantity of ‘voweled stem canonic patterns’ for action words and nouns is almost 10,000 [6]. Some problems in Arabic texts include considerable translations and translated labeled entity, its satire is usually contradicted texts on Arabic. Despite the fact that the methodology demonstrated to create high exactness, it has a few flaws. It requires building a tremendous corpus (dataset) and marking it manually by human experts. The system of manual commentary can be extremely troublesome in any event, for native speakers because of criticism and social references. It can also be costly and tedious. A constraint is that NLP tools designed for Western dialects are not effectively versatile to Arabic due to the specific highlights of the Arabic language [7].

2 Related Work

Many different researchers have focused on part of speech tagging in many languages like Hindi, English, Arabic, Chinese, Marathi, and others. Here some important related work in part of speech tagging techniques is discussed.

Singa et.al. (2012) has advanced part of speech tagger on based rule approach with a rate of accuracy 50, 77, 85% on lexicon data words (50-100-1000) in Manipuri language (Manipuri is mightily Speaker in Manipur, Bangladesh, Tripura, and Myanmar, Assam [8].

Zelalem (2013) has used a hybrid approach of HMM and rule-based tagger, on a collected dataset of 354 sentences with accuracy of 77.19, 61.88, and 80.47% in Kafi- **Noonoo** language (language in southwestern Ethiopia) [9].

Deepali et al. (2018) has designed the POS tagging base rule in the Marathi language, on a collection of 1364 words with an accuracy rate was 100% [10].

Aliwy et al. (2018) suggested a new approach using HMM and n-grams taggers for tagging Arabic words in a long sentence to collect 1000 datasets as documents and 526,321 as a separate token, this system gives an accuracy rate of 0.888, 0.925, and 0.957 [11].

Barud et al. (2019) developed parts of speech tagger for Awngi language using Hidden Markov Model (HMM), in a dataset 94,000 sentences were collected, total word of 188,760 and attained accuracy rate of 93.64 and 94.77% [12].

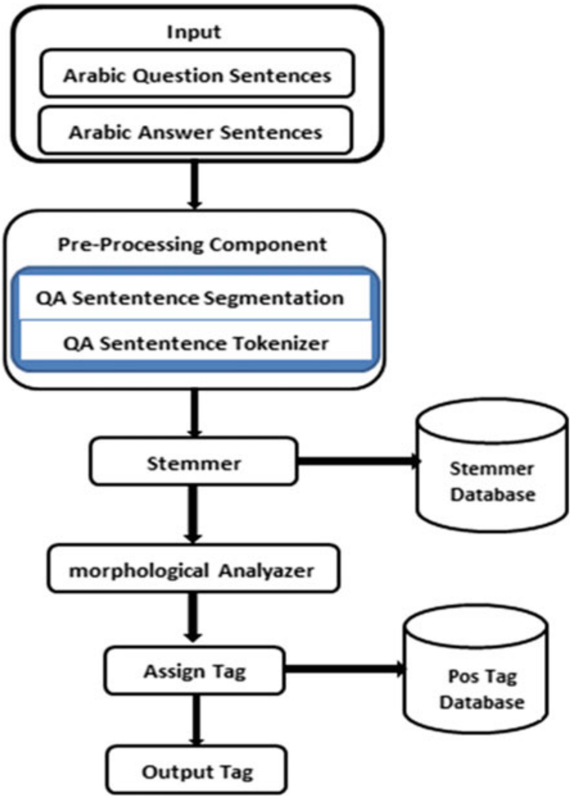
Hagos, 2020 designed a Ge’ez (Ethiopian language) POS tagging using a Hybrid approach, the data set collected was 15,154 words, 1,305 sentences, and the result accuracy rate was 77.87, 82.23, and 94.32% [13].A detailed description of building a POS tagger is discussed in further sections.

3 Proposed System

3.1 Data Collection

The proposed system considers 40 questions and 40 answers that are taken from five different lessons of the 4th standard textbook of the Arabic language as the raw input. The data is pre-processed, and a well-designed stemmer is developed for supporting the morphological analyzer. An output of a morphological analyzer is used as an input to the part of speech tagging. A linguistic rule-based POS tagger is developed, and ultimately, the tag data is provided by the system. This tagged data will be used in post-processing for better performance of the QA system. The detailed block diagram of the system proposed is shown in Fig. 1, and further stages have been discussed in detail with example.

Fig. 1 Proposed system



3.2 Pre-processing Component

3.2.1 Questions and Answers Sentence Segmentation

In the Arabic language, the structure is dissimilar for different sentence types. The proposed system consists of two segments namely, question and answer. The former is the question part, and the latter is the answer part. The input data is taken from the 4th standard textbook of Arabic language. The questions are framed using the textbook lesson and the number of word used are confined. The answers for the question are stored in the answer file and the answers are based on the lessons.

3.2.2 Questions and Answers Sentence Tokenization

Tokenization is the method of tokenizing or parting a string, text into a listing of tokens. One can consider token parts like a word is a token in a sentence, and a sentence is a token in a passage. This process of separating tokens is the input text.

Each word is separated from the sentence considering white space or symbols as one token and treat each word individually. Then POS needs to split the input text into tokens tagging. The tokenization of the question sentence is shown below:

' لماذا يعد البن اليمني من اجود انواع البن في العالم ؟ '

and word tokenization has split answer sentence into words as follows:

' يعد' ' البن' ' اليمني' ' من' ' اجود' ' انواع' ' البن' ' في' ' العالم' ' لما' ' يمتاز' ' بة
'' من' ' نكهة' ' طيبة ' '.'

3.3 Stemming

Stemming is the operation of decreasing a word to its word stem that contains affixes to suffixes and prefixes or the roots of words known as the lemma. A rule-based steamer uses specific pre-defined rules according to language to mark another type of word used in its base. These language connected rules are created manually by language practitioners. Rule-based stemming methods are divided into three categories such as morphological, table lookup, and affix stripping. Arabic stemmer is a very different and difficult structure than other languages. Stemming is very necessary for natural language understanding and natural language processing. Arabic word language structure is a grammatical rule on the root, and pattern scheme, its deliberate as a root based language with more than 10,000 roots. Arabic words are commonly founded on three-dimensional roots: three consonants, which characterize the hidden significance of the word. Diverse long and short vowels, prefixes, and postfixes are added to that root to make the ultimate desired inflection of sense. These modifications follow designs that reflect across roots. Stemming is the only source to extract the root in the Arabic language. For example, the Arabic word ' المعلمات 'contains the following component in Tables 1 and 2.

Figure 2 depicts an example for code execution and the flow of stemming in the proposed model (Figs. 3 and 4).

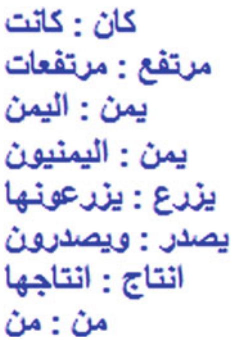
3.4 Morphological Analyzer

The morphological analyzer of Arabic words is a procedure for each word of the input text to choose its root and pattern. The outcomes of the morphological analyzer can be used for further analysis[14]. The morphological analysis goals are to train the internal structure of a word. Words after molding are analyzed to check if they are sorted or not. When a stem word is produced, then the word root is formed

Table 1 Example of Arabic Affix

Word	root	prefix	Suffix	Infix
المعلمات	علم	ال	ات	ا

Fig. 2 Result for Arabic stemming from QA Sentences



كان : كانت
مرتفع : مرتفعات
يمن : اليمن
يمن : اليمنيون
يزرع : يزرعونها
يصدر : ويصدرون
انتاج : انتاجها
من : من

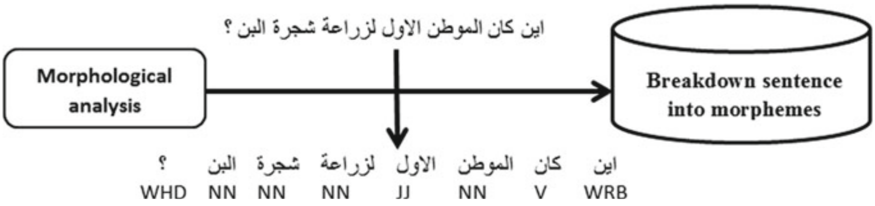
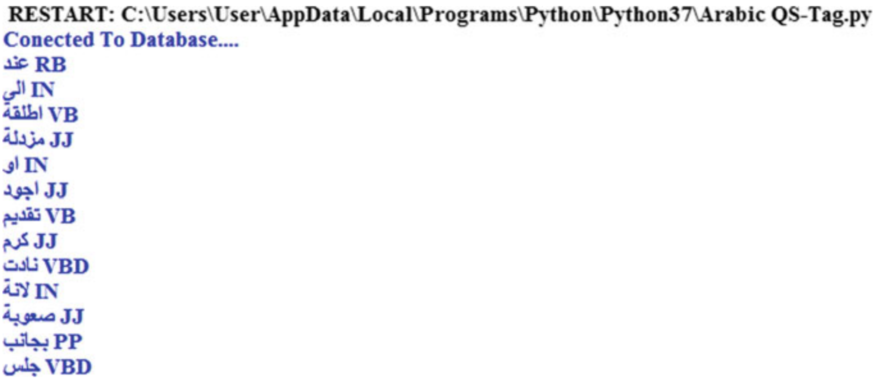


Fig. 3 Role of the morphological analysis of Arabic words



```
RESTART: C:\Users\User\AppData\Local\Programs\Python\Python37\Arabic QS-Tag.py
Conected To Database....
عند RB
الي IN
اطلقة VB
مزدلة JJ
او IN
اجود JJ
تقديم VB
كرم JJ
نادت VBD
لانة IN
صعوبة JJ
بجانب PP
جلس VBD
```

Fig. 4 Result of rule-based POS tagger for QA in the Arabic language

by combining substituted letters with the stem word. The morphological analyst is expected to produce root names for the given input document [2].

Table 2 Arabic prefixes

Words Example	Prefix
بالحديقة	ب
كدواء	ك
والاولاد	و
الزراعة	ال

Table 3 Rule-based POS tagging and Farasa POS tagging

Rule-based POS tagging	Farasa POS tagging	Word	NO
RB	NOUN	عند	1
IN	NOUN	الى	2
VB	NOUN	اطلقة	3
JJ	NOUN	مزدلة	4
IN	NOUN	او	5
JJ	NOUN	اجود	6
VB	NOUN	تقديم	7
JJ	NOUN	كرم	8
VBD	NOUN	نادت	9
IN	NOUN	لانة	10
JJ	NOUN	صعوبة	11
PP	NOUN	بجانب	12

3.5 Tag Generation

Initially, Farasa is attempted to obtain the tagged word, but the results were not that much appreciable. Table 3 shows the incorrect tags using Farasa.

A linguistic rule-based POS tagger is chosen to be developed due to the limitations over the Farasa.

3.6 Rule Based POS Tagger

The data is manually analyzed to find the context for each part of speech and based on the data a rule is developed.

(a) **Algorithm for POS Tagging System**

Step 1 Input the question and answer sentence segmentation.

Table 4 Proposed system data

Lessons	No. of (Q & A)	No. of words (Q & A)
L1	16	134
L2	16	100
L3	16	123
L4	16	135
L5	16	131
Total Words		623

Step 2 Start to tokenize the Q & A sentence into words.

Step 3 Generate a stemmer for all words to obtain the original words by using a morphological analyzer.

Step 4 Gather all the incorrect tags for all the words from the Farasa parser as a word by word and transfer to the database to produce the correct tag for all words.

Step 5 Allocate a suitable tag to append to its word.

(b) **Result**

In this article, the data were collected from the Arabic language among one of the subjects from the 4th standard Arabic medium state board pattern in Yemen. Five different lessons are considered in this work and formed 40 questions and 40 answers respectively. Table 4 illustrates the data in lessons.

The formula to calculate the accuracy and the performance in Farasa parser tagging and rule-based POS tagging is applied:

Accuracy = (No. of Correctly tagged/ Total No. Tagged in documents)*100.

Subsequently, the accuracy result in Farasa parser was 92.77%. and the proposed system accuracy is 100%.

4 Conclusion and Future Scope

Question Answering System comes beneath the purview of natural language processing, thus natural language understanding is an essential task. For the present work, the corpus (database) for the QAS is build based on the Arabic language subject of the 4th standard Arabic medium state board pattern in Yemen. POS tagger is a very important component of any QA system as it impacts the accuracy of question answers generated for the system. In thisarticle, the major inputs to the POS tagger is based on sentence splitting, word tokenization, and stemming which serves as the basis for better performance of POS tagging. Linguistic POS tagger rule-based design and implementation are presented in detail concerning the QA system. The results are compared with Farasa and determined that our rule-based system yielded better results. In the future scope, this tagged data will be employed to measure the results of the QA system.

References

1. Shaalan K, Siddiqui S, Alkhatib M, Monem AA Challenges in Arabic natural language processing , School of Informatics, University of Edinburgh1, UKFaculty of Computer and Information Sciences, Ain Shams University 4, Abbassia, 11566 Cairo, Egypt
2. Yao X (2014) Feature-driven question answering with natural language alignment , Johns Hopkins University, Doctor of Philosophy thesis 2014
3. AL-Taani A, Abu Al-Rub S (2009) A rule-based approach for tagging non-vocalized Arabic words. Int Arab J Inf Technol 6(3)
4. Habash NY (2010) Introduction to Arabic natural language processing, A Publication in the Morgan and Claypool Publishers series. ISBN: 9781598297966
5. https://arabicquick.com/an-introduction-to-the-arabic-language
6. Darwish (2002) Building a Shallow Arabic morphological analyzer in one day. In: Proceedings of the ACL workshop on computational approaches to semitic languages, Philadelphia, PA, pp 1–8
7. Ezzeld AM, Shaheen M (2012) Survery of Arabic question answering: challenges ,tasks , approaches , tools , and future trends. In: The 13th international Arab conference on Information technology ACT2012. ISSN.1812–0857
8. Raju Singha KH, Purkayastha BS, Singha KD (2012) Part of speech tagging in Manipuri: a rule-based approach. Int J Comput Appl 51(14): 0975–8887
9. Mekuria Z (2013) Design and development of part-of-speech tagger for Kafi-noonoo language, Master's thesis, Addis Ababa University, Addis Ababa
10. Deepali G, Naik Ramesh R, Namrata Mahender C (2018) Rule-based part-of-speech tagger for Marathi language, © 2018 IJSRST | vol 4, issue 5. Print ISSN: 2395–6011. Online ISSN: 2395–602X Themed Section: Science and Technology
11. Aliwy AH, Al_Raza DA (2018) Part of speech tagging in Arabic long sentence. Int J Eng Technol 7(3.27):125–128
12. Wubetu Barud Demi lie, Parts of Speech Tagger for Awngi Language. ISSN 2321 3361 © 2019 IJESC
13. Gebremedhin H, gebremeskel S (2020) Ge'ez POS tagger using hybrid approach. Int J Comput Sci Inf Technol Res 8(1):12–23
14. Awajan A (2018) A rule-based morphological analyzer of Arabic words. https://www.researchgate.net/publication/330006249

Change Detection in Land Use-Land Cover Using Convolutional Neural Network

Sahithi Samudrala, Mekala Pranvitha nand, Suhail Mohammad, and Radhesyam Vaddi

Abstract Change detection in satellite imagery is a concerning topic for researchers and scientists. Studying changes in satellite images gives an enormous amount of knowledge in studying the geography and ecosystem. The dataset used in this work is taken from the AVIRIS sensor and normalized to reduce the time of execution. This paper covers the important aspects of the classification of satellite imagery using convolutional neural networks. Changes are detected based on the classification maps in multi-temporal imagery. Experimental results show that the proposed method got accuracy on par with state-of the-art methods.

Keywords Satellite imagery · Convolutional neural networks

1 Introduction

The changes on the earth's surface from natural and human causes are rapid. Scientists and researchers are concerned about detecting the changes without knowing the scale of changes which is very hard and need to be addressed [1]. Many techniques had been developed for detecting changes on land cover which are optimal for both satellite imagery and also for digital imagery [2]. Satellite images are different from normal images. They are not to be compared with each other. We can adapt the techniques or algorithms used for normal imagery on satellite imagery. Normal images are three or four channels and are speeded over the visual spectrum. Satellite images are multiple channels and are ranged from infrared to ultraviolet in the electromagnetic spectrum

---

S. Samudrala (✉) · M. P. nand · S. Mohammad · R. Vaddi
Department of Information Technology V.R, Siddhartha Engineering College, Vijayawada, India
e-mail: sahithi.rocking25@gmail.com

M. P. nand
e-mail: mekalapranvithanand@gmail.com

S. Mohammad
e-mail: 8125527706ms@gmail.com

R. Vaddi
e-mail: syam.radhe@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_59

with minimal intervals. Digital images represent the intensity of the red–green–blue in the image, whereas satellite images are reflection indexes for each interval [3].

The push-broom passively was considered for scanned images [4]. The spectral images can be collected in multiple ways from satellite to airborne imaging through more complex spectroscopes. Spectroscopes scan land cover in the spatial perspective in all the intervals of the electromagnetic spectrum collecting spectral data.

The spectral data collected for each spatial pixel is labeled manually by surveying onsite, and respected ground truths of the spectral data are provided [5]. But it is not possible to survey land in a remote region manually, so we are using a deep neural network to learn the patterns of each spectral data in a remote region from the knowledge of spectral data which was provided with ground truths collected by manual surveys generate the classification maps aka ground truths [6].

The full article is organized as follows. Section II describes the project that needs to be changed for the detection of the dataset. Section III explains the CNN algorithm and Section IV depicts the methodology for the proposed system. Section V presents a complete analysis of the results. Finally, Section VI describes the conclusion and future work.

2 Dataset Description and Normalization

The dataset used in this work is from the AVIRIS sensor [7] with the following features.

(a) Each class is uniquely represented by an integer in the ground truth.
(b) They are 16 classes in the dataset
(c) The spatial resolution of the imagery is 145 × 145 pixels with a spectral resolution of 220 channels for each pixel.

The data consists of a wide range of variables with very high numeric values. Cleaning the outliner values and normalizing the entire data was done for reducing the computation stress. Reducing the scale can fasten the computation, and there is no loss in the pattern. The patterns are shown in Fig. 1 (Fig. 2).

3 Algorithm

There is a wide choice of selection in deep neural networks for supervised learning. The primary reason is to choose a convolutional neural network and its unique feature to detect new patterns. Recent studies prove that this is one of the best approaches in supervised learning mechanisms.

(A) *There are 12-layered convolutional neural networks (CNNs), three convolutional layers, and three max pooling layers that are accompanied by a fully* *connected neural network layer.* Figure 1*explains the individual layers—convolutional layer.*

There are three convolutional layers used with ReLU as activation function and with filters of size (200,100,100). This is using with a kernel size of (2,2) for the first conv-layer and (1,1) for the next two layers.

(B) *Pooling layer*

(C) *The max pooling layer is used with a kernel size of (2,2), and each layer of pooling uses the ReLU as an activation function—fully connected layer.*

The fully connected layer has four dense layers; the input for the first fully connected layer was flatten, and then return to the first layer, and each layer has ReLU as the activation function except the last output layer which delivers activation as softmax, which returns the probability in the range of zero to one.

4 Methodology

The change is detected in between two images which are taken from a satellite image [8]. Two land-cover images are taken from different timestamps 2013 and 2015. Figure 3 represents the methodology of the proposed model (Fig. 4).

(A) *The two classification maps such as classification map1*map *2 are obtained using the CNN classifier. A simple difference operation is performed on these two maps and recognized the change map with 97% accuracy* [9].*Training*

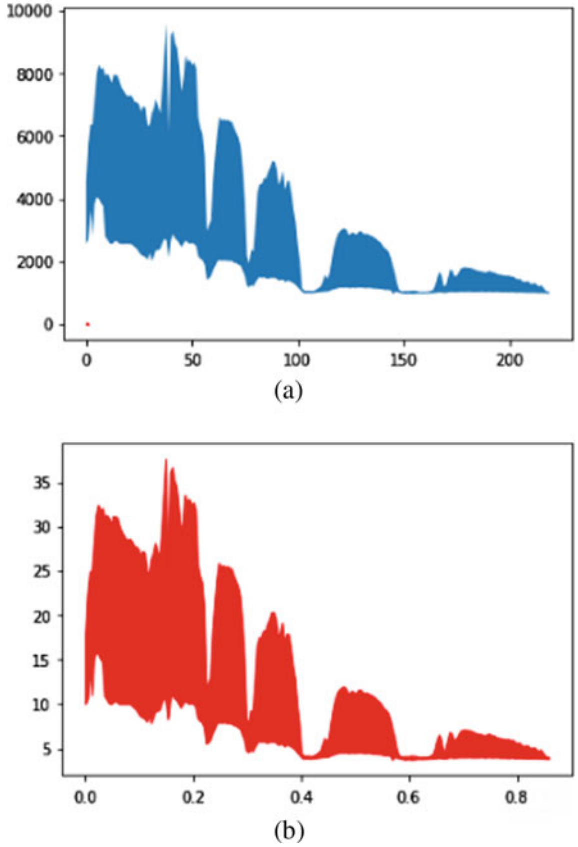
The training was performed on 8199 samples, and the test size was 2050. The model was trained for 20 epochs with a learning rate of 0.01. It took 14 s to complete each iteration, i.e., an average time of 5 min to get a training accuracy of 99% and a test accuracy of 99.66%. The learning curve is shown in Fig. 5 and the loss curve in Fig. 6.

5 Results

Figure 8 describes the results and the change map obtained from the two images that are shown in Fig. 7. Consider these two images they are inputs taken through satellite using Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor data which is temporally varied images that means considered in two different timestamps that are 2013 and 2015 from the area Hermiston located in Oregon United States. The width, height, and band of the channel (or) channel color are (390,240,242), respectively.

Both the images are classified using a classifier that is by adopting the convolutional neural networks (CNN) algorithm which produces two classification maps by undergoing the process of CNN; they are conv-layer, batch normalization, max pooling, dropout, flatten, and dense. Figure 7a, b describes the CNN process and the

Fig. 1 Scale graph of the data **a** before normalization **b** after normalization



(a)

(b)

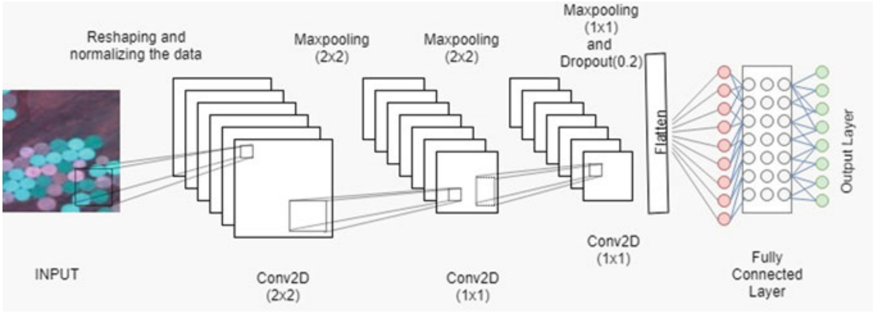


Fig. 2 Different layers of the convolutional neural network

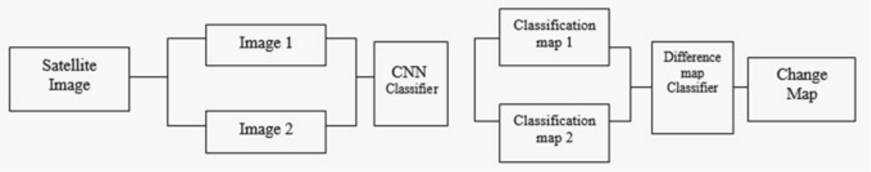


Fig. 3 Architecture diagram of the proposed change detection method

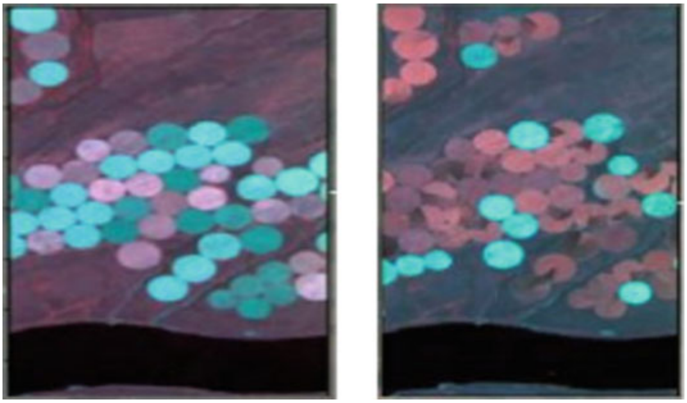


Fig. 4 Satellite images used for change detection

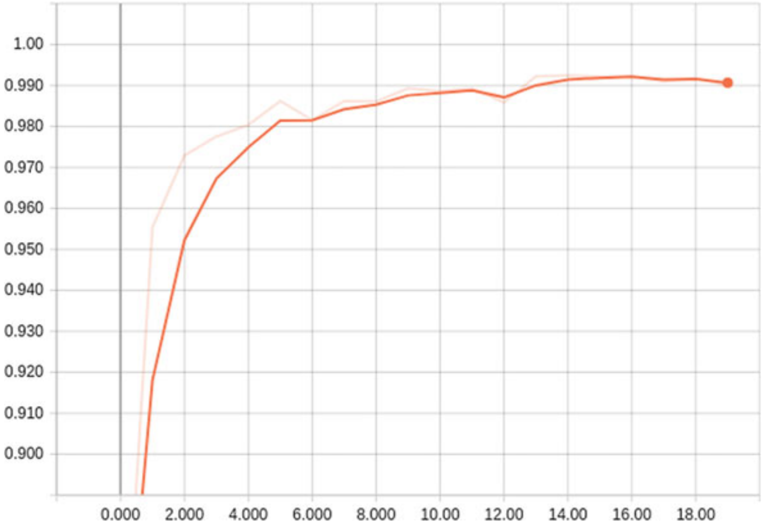


Fig. 5 Learning curve

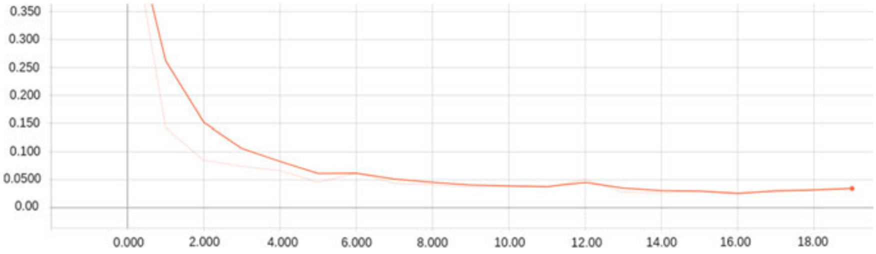


Fig. 6 Loss curve

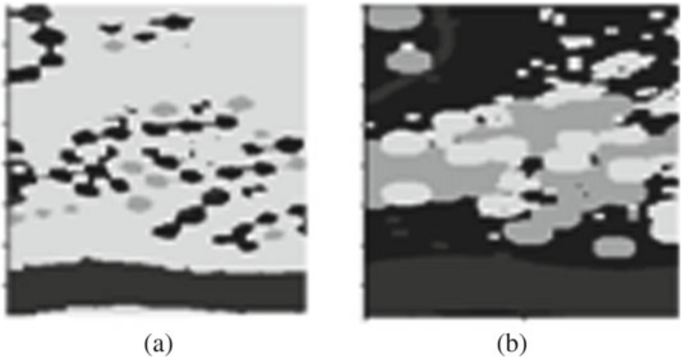
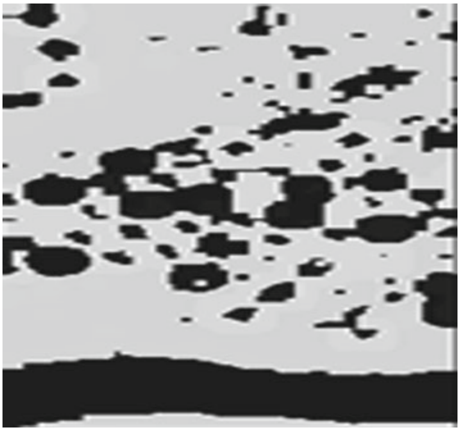


Fig. 7 Classification maps **a** 2013 **b** 2015

Fig. 8 Change map



two classification maps. The height and width of these images will be (390,200); no band channel is analyzed.

To generate the change map results, it is necessary to find the difference between the two classification maps taken from two different timestamps which are shown in

Table 1 Changes notified with number of pixels table-type styles

Class	Year 2013	Year 2015	Difference
1	31,456	8269	23,187
2	14,967	11,122	3845
7	16,239	2236	14,003
9	15,338	5373	−41,035

Fig. 7. In the figure, the yellow area indicates the changed area and purple indicates the unchanged area.

The following Table 1 illustrates the changes in the classes. The first column specifies the classes, and the remaining columns are the exact count of class labels with difference in the last column.

6 Conclusion and Future scope

The technique of using a deep convolutional neural network model for generating classification maps is optimal for this use case. The generated classification maps for two temporal images are parsed for generating change maps. This kind of approach is most beneficial for detecting the changes in the known classes. And this method has been widely used in applications such as agriculture, forestry, and mineralogy.

The change map obtained was strong enough to identify the total changes in the complete spatial resolution. The class-wise changes are attained from the convolution neural networks output layer. Further, the study can be made in developing a hybrid convolutional neural network, which can classify two-satellite imagery and identify the difference in the classification maps by using the dual pipeline method.

References

1. Lu D, Mausel P, Brondizioand E, Moran E (2004) Change detection techniques. Int J Remote Sens 25:2365–2407
2. Coppin PR, Bauer ME (1996) Digital change detection in forest ecosystems with remote sensing imagery. Remote Sens Rev 13:207–234
3. Singh A (1989) Digital change detection techniques using remotely-sensed data. Int J Remote Sens 10:898–1003
4. Blaschke T Towards a framework for change detection based on image objects. Göttinger GeographischeAbhandlungen, p 113
5. Zhang H, Gong M, Zhang P, Su L, Shi J (2016) Feature-level change detection using deep representation and feature change analysis for multispectral imagery. IEEE Geosci Remote Sens Lett 13
6. Doña C (2014) Empirical relationships for monitoring water quality of lakes and reservoirs through multispectral images. IEEE J Select Topics Appl Earth Observ Remote Sens 7(5)

7. Baumgardner MF, Biehl LL, Landgrebe DA (2015) 220 Band AVIRIS Hyperspectral Image Data Set: June 12, 1992 Indian Pine Test Site 3. Purdue University Research Repository. https://doi.org/10.4231/R7RX991C
8. Huang C (2016) Surface water change detection using change vector analysis IEEE
9. Gong M,Zhan T, Zhang P, Miao Q (2017) Superpixel-based difference representation learning for change detection in multispectral remote sensing images. IEEE Trans Geosci Remote Sens

An Efficient Ant Colony-Based Self-adaptive Routing Strategy for IoT Network Model

K. M. Baalamurugan, D. Vinotha, and Premkumar Sivakumar

Abstract The intellectual device is known as the Internet of Things (IoT) not only offer services, but also facilitate the allocation of heterogeneous resources and diminish resource utilization based on service time. This is considered as a crucial fact toward the large-scale environments. To make the service effectual with the proper response, the request for the route has to be analyzed simultaneously and to offer better global solutions. Therefore, this work anticipates a self-adaptive ant colony optimizer (SA-ACO) for IoT environment. This SA-ACO model relies on the provisioning of a global solution with the self-adaptive characteristics and modeled to select the finest attributes. The node clustering is dependent on the population and to provide the functionality over multi-directional and the resourceful realization of self-adaptive route searching. The memory utilization is analyzed with the selection of nominal solutions to acquire promising results. To validate the competency of the anticipated model and to measure the effectiveness, this work can be analyzed over the real-time environment in future. The simulation is performed in the MATLAB environment that the anticipated model can acquire the finest solution with strong exploration functionality, and superior performance is evaluated with the comparison of the model with PSO and RR, respectively.

Keywords Internet of Things (IoT) · Self-adaptive ant colony optimizer (SA-ACO) · Artificial intelligence (AI)

1 Introduction

The advancements of the Internet of Things (IoT) have entered a progression of intelligence industries. Recently, the energy savings, configuration optimization,

---

K. M. Baalamurugan (✉)
School of Computing Science and Engineering, Galgotias University, Greater Noida, UP, India
e-mail: k.baalamurugan@galgotiasuniversity.edu.in

D. Vinotha · P. Sivakumar
Department of Computer Science and Engineering, Annamalai University, Chidambaram, TN, India
e-mail: vinothacse54@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_60

and environmental protection toward IoT resources have turn toward diverse issues that has to be resolved [1]. There are diverse artificial intelligence approaches for resource scheduling that are modeled for a certain application and not applicable for resolving IoT service crisis examined [2]. For provisioning the service, this is based on the complete IoT layout for service system. The resource optimization allocation problem based on service-oriented network collaborative model is extremely complex [3]. This comes under the typical NP-hard combinational optimization crisis. The fact is that how to reduce the resource utilization, and reduce the response time. Subsequently, how to multi-optimal services in numerous candidate set is chosen to fulfill the objectives [4]. Therefore, this is a challenging factor in multi-objective optimization crisis.

Various investigations have tried to resolve the multi-objective service selection problem in internet services. The author in [5], depicted that the approximation strategy for multi-objective driven service selection. The author in [6], provided research on multi-objective optimization to reach the QoS. This model anticipated a parent set model for QoS aware service composition. The author in [7], anticipated a model that assists decision making in predicting effectual, QoS optimization service with clustering. The prevailing research model is executed with an adaptive service composition acquired from the immune system. Moreover, the above-mentioned model is concentrated base don QoS.

The features with IoT services are heterogeneity, large-scale, dynamic, and unreliability in nature that shows differentiation in web services. An effectual confront to resolve the IoT service domain is the construction of effectual service selection procedures for optimal management of both QoS and energy. This crisis is crucial over large-scale IoT environments that comprise of a huge amount of distributed entities. The author in [8] depicts that IoT is a paradigm where real-world physical things that are associated with the services and internet provisioning via the computational devices. The three QoS scheduling approach for service-based IoT was anticipated by Sherubha and Mohanasundaram [9]. The sensing-based service model is considered to be constructed on the top of IoT services and infrastructure. The author in [10], allocated services to interface heterogeneous resources and generates an optimal solution for computation for a hard problem.

Moreover, to determine the coordination among the environment and population, some population are based on evolutionary procedure, co-evolutional methods are initiated toward the immune-optimization model, and superior outcomes are attained for resolving the combinatorial optimization crisis. The cooperative and competition model are two essential factors over the multi-objective optimization crisis. This model relies on the attainment of higher success in resolving single-objective optimization issues. The author in [11], provided an evolutionary algorithm for a multi-objective process which is competent in preserving the archive diversity by the huge operator and dynamic sharing. The author anticipated competitive evolutionary model. There are diverse sub-populations that are optimized for decision variables, respectively. The differences among the models are mapped in association with every sub-population and decision variables that are not centralized; however, it is demonstrated with the superior outcomes.

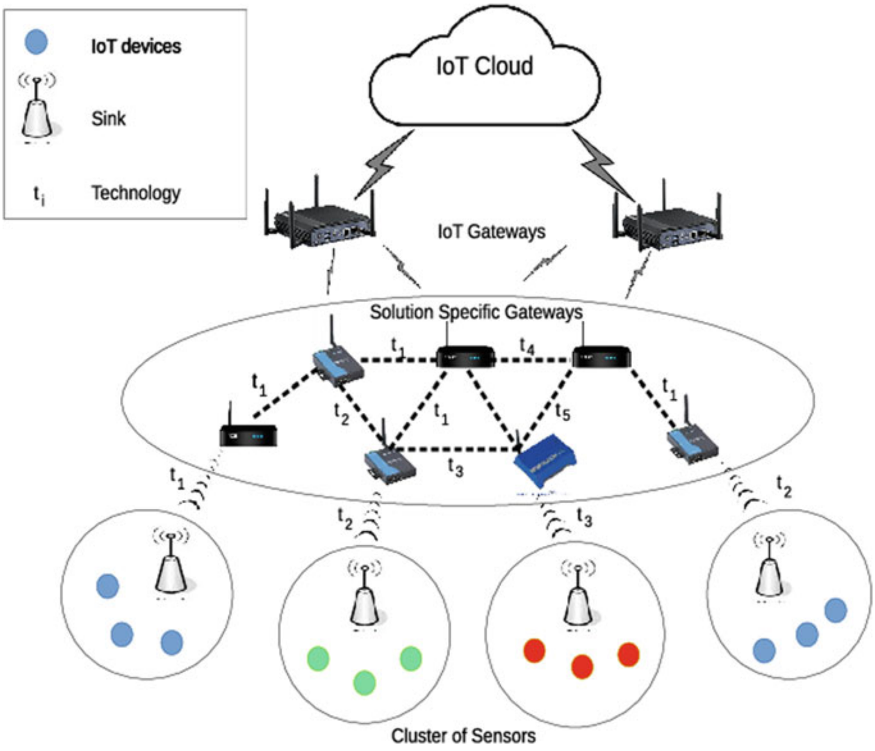


Fig. 1 Generic view of IoT-based routing model (*Source* Angelakis et al. [2])

The significant contributions of this work are:

1. With the inspiration from the meta-heuristic model, an adaptive approach is modeled and designed with sub-population to attain a global solution that assists effectual interactions between the populations. This is also based on global optimization solutions.
2. The self-adaptive nature of ant to gather the food and to reach the source is examined and the evolution of ant population, that is resourcefully resolved with the selection of non-dominated outcomes.
3. The statistical and the clustering model with the heuristic process, varying components are examined, that produces the operations more purposefully and directionally.

The sections of the paper are as follow: Sect. 2 depicts the background model of prevailing approaches, Sect. 3 is the anticipated self-adaptive ant colony optimization method for routing. Section 4 shows the numerical results and the validation, and Sect. 5 draws the conclusion and future research directions.

2 Related Works

Recently, evolutionary algorithms are turned to be a main-stream model for performing various researches in the soft computational field. The author in [11] provided a comprehensive analysis of the modern multi-objective evolutionary process. However, the multi-objective optimization process relies on the artificial immune system that comprises of multi-objective problem, constrained problem, artificial network model for the vector-based immune system, non-dominant immune model, and so on. However, based on the analysis, the author in [11] anticipated a genetic model and provides various strategies for choosing the vaccines and develop the immune operators. The author in [12] provided a biological notion with the vaccine usage for promoting exploration in search space.

In recent times, a novel immune model for MOO is anticipated by Sherubha and Mohanasundaram [12]. Here, the de-generation for predicting the clonal selection process for multi-modal optimization is examined. The novel multi-class clustering approach is based on the maximal margin-based clustering approach, and an evolutionary model is anticipated. However, this model concentrates on the local searching process. The local improvement operator with (convergence acceleration operator) was initiated, and the hill-climbing with sidestep is modeled for local searching, where a novel multi-objective optimization structure relies on non-dominated searching and local sorting was anticipated. The author [13] used a novel ranking approach termed as global margin ranking with position deployment of every individual in objective space to attain marginal dominance all through the population.

In [14], the anticipated novel co-evolutionary model that relies on an elite approach, where every elite individual are utilized for guiding the search function. The author in [14], modeled a structure termed as hyper-MOO. The sub-population size was modified based on the appropriate MOO performance. The author used multiple sub-population and utilized diverse statistical and clustering techniques to assist the new population generation and local searching. However, certain investigators concentrate on collaboratively decomposing multi-objective optimization issues. Additionally, it is inspired by the mammalian endocrine system with an artificial endocrine model for managing power that is modeled with robotic analysis. This mechanism is anticipated to control the cooperative evolution among the particles.

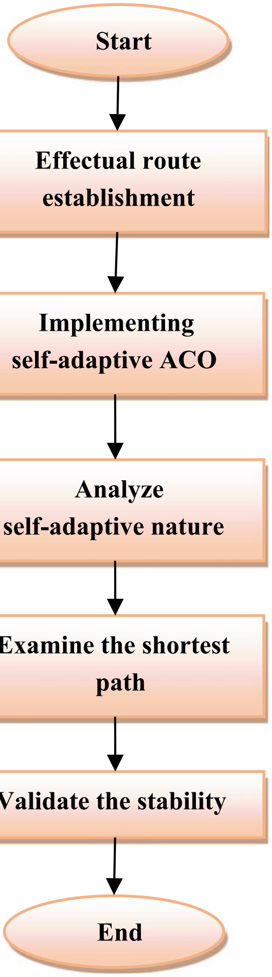
In [15], has utilized the evolutionary model and diverse populations for multi-objective optimization issues. With the immune-based model, there are diverse sub-population are evolved with diverse evolutionary approaches. The immune co-evolutionary model possess two phases that are modeled to acquire the optimal partitioning balance. The author in [16], provided an evolutionary model for dealing with the garmenting problem that initiates the distance affinity and dominance affinity. The author anticipated evolutionary methods for handling routing issues. The investigator used two diverse sub-populations that are co-operatively evolved with the evolution process to attain superior global optimality for radial basis function estimation with neural networks.

3 Methodology

This section discusses the self-adaptive ant colony optimization approach used for handling effectual routing over IoT environment. The flow diagram of SA-ACO is given in Fig. 2.

In general, ACO is a popular meta-heuristic approach that is inspired by the ant's ability to determine the shortest path among the food sources and the colonies. This foraging nature for collecting the food source is utilized to resolve the diverse optimization issues like the prediction of the shortest path to find the corresponding neighborhood for collecting the food. The ants pretend to communicate with one another for realizing the chemical signals termed as pheromones. While moving back to their location, the ant releases a certain amount of pheromone with path proportional to the

Fig. 2 Flow diagram of self-adaptive ACO


Start
Effectual route establishment
Implementing self-adaptive ACO
Analyze self-adaptive nature
Examine the shortest path
Validate the stability
End

quality and quantity of sources. However, all the other ants possess superior probability for the successive paths with higher pheromone concentrations. Moreover, all ants have to follow the shortest path for the source to the home. Consider, 'm' ants and 'n' probable paths where every ant chooses the probable path based on the highest concentration over the pheromones with probable paths.

The anticipated self-adaptive ACO approach is considered to be an effectual model with heuristic nature for various NP-hard-based research issues that includes job-shop scheduling and traveling salesman problems. Moreover, the anticipated ACO model is used for scheduling issues over the cloud and includes scheduling over the VM on the cloud environment and the scheduling tasks for the VMs to attain the load balancing tasks in VM and reduces the response time during task functionality. Also, the ACO algorithm is utilized for scheduling various IoT tasks over the cloud. Added, this algorithm is utilized for providing the deadline-aware task scheduling process over the IoT-based computational infrastructure. The anticipated model concentrates in improving the profits of service providers while determining the fulfillment of deadline-based IoT task constraints.

While performing the route identification tasks with the help of connected nodes with an ultimate target to reduce the response time, the ant selected the node for improving the stability during task assignment with the probability which is expressed as in Eq. (1).

$$P_{ij}^{k}(t) = \frac{\left(\tau_{ij}(t)\right)^{\alpha}\left(\eta_{ij}(t)\right)^{\beta}}{\sum_{s}(\tau_{is}(t))^{\alpha}(\eta_{is}(t))^{\beta}} \tag{1}$$

Here, α and β are heuristic constants, $\alpha \geq 0$ is a heuristic factor that manages the consequences of pheromone quantity with $\beta \geq 1$ is a heuristic parameter that deals with the significance of route allocation. $\eta_{ij}(t)$ is a heuristic model that specifies the routing stability and is computed with Eq. (2).

$$\eta_{ij}(t) = \frac{\text{load}_j}{R_{ij}} \tag{2}$$

Here, R_{ij} is evaluated and load_jspecifies the load over the node and computed with the equation mentioned above. However, R_j increases when load_jreduces and $\eta_{ij}(t)$ reduces. As an outcome, the routing stability of the nodes over the network is evaluated. $\tau_{ij}^{k}(t)$ is pheromone trail quantity with allocating the task to nodes by route prediction with t time and $\tau_{ij}^{k}(t+1)$ is pheromone trail for allocating tasks for all ants during the iteration $t+1$ which is evaluated as in Eq. (3).

$$\tau_{ij}^{k}(t+1) = (1-\rho)\tau_{ij}^{k}(t) + \rho\Delta\tau_{ij}^{k}(t) \tag{3}$$

Here, $\Delta\tau_{ij}^{k}(t) = \frac{1}{R_{ij}}$ and ρ is constant that specifies the pheromone evaporation rate that stimulates the evaporation effect of pheromones during every step. In addition, the pheromone trail is updated globally with all ants with probable task determination over the complete iteration process. This global updation is expressed as in Eq. (4).

$$\tau_{ij}^{k}(t+1) = \left(1 - \rho_g\right)\tau_{ij}^{k}(t) + \rho_g \Delta\tau_{ij}^{k}(t) \tag{4}$$

Here, $\Delta\tau_{ij}^{k}(t) = \frac{1}{L_{\text{best}}}$, L_{best} is the finest route identified with ρ_g global evaporation rate. The initial pheromone value is computed with Eq. (5).

$$\tau_{ij}^{k}(0) = \frac{R_{ij}}{R_{\text{Average}}} \tag{5}$$

Algorithm

```
Parameter initialization
Compute $R_{ij}$
Initialize pheromone matrix
Compute $\eta_{ij}^{k}(t)$
While $iteration \leq N_{iteration}$ do
    Position the ant to initiate the route randomly
    For $ant = 1toN_k$ ant = 1toN$_k$ do
        For $S_i = 1$ s$_i$ = 1 to nodes
            Compute $P_{ij}$
            Add the stability route to the routing table
        End for
        Revise $\tau_{ij}^{k}(t+1)$
    End for
    Evaluate the prior solution and update the best]
    If solution is best then
        Revise $\tau_{ij}^{k}(t+1)$
    End if
    $iteration = iteration + 1$;
endwhileend while
Clear routing table
```

The anticipated SA-ACO model is utilized to determine the effectual process of measuring route stability over the IoT environment. The nodes have to fulfill the QoS constraints with appropriate response time by determining its self-adaptive nature. The load and the service time of the nodes are measured with the initialized parameters. The heuristic model determines the maximal amount of iterations and initializes the total amount of nodes over the network with the productivity rate of all sensors and service rates.

The successive steps are applied for measuring the response time of nodes with the provided equations. The initial pheromone and task allocation stability for all nodes are given in the equation above. With all iteration, the ant has to offload the computational measure of the nodes with probability evaluation. The node selection is done with the clustering method. This method evaluates the solution probability and the cumulative probabilities that are assessed in an ascending manner. The random numbers are produced with the cumulative probability. At last, the appropriate solutions are generated with a random number. This is validated as the effectual selection of resourceful solutions. When the ant completes the routing process over the connected nodes, the local pheromone is updated. With all iterations, the algorithm needs to revise the global pheromone. The iteration is completed until the maximal amount of iterations are attained.

4 Numerical Results

The MATLAB simulation environment is run over PC with Intel core Pentium process with 2.60 and 8 GB RAM. With this experimentation, various parameter settings are given below. The parameter settings are provided by performing numerous primary experimentations.

The experimentation is performed to measure the anticipated model's ability for stabilized routing establishment with average response time, routing imbalance, and node connectivity, respectively. The imbalance degree is expressed as in Eq. (6).

$$\text{Degree of imbalance} = \frac{Max\left(R_j\right) - Min\left(R_j\right)}{R_{\text{Average}}} \tag{6}$$

The response time-based SD is computed with the route distribution to establish the stability, which is expressed as in Eq. (7).

$$\text{SD} = \sqrt{\frac{\sum_j \left(R_j - R_{\text{average}}\right)^2}{N_{\text{nodes}}}} \tag{7}$$

From Tables 1, 2 and 3, it is observed that the metrics like nodes imbalance degree, average node connectivity, and route response are measured, respectively. The anticipated model is evaluated and compared with PSO and RR where the performance of SA-ACO is superior in contrast to other models. Similarly, Figs. 3, 4 and 5 depicts the graphical representation of the anticipated approach. This model works effectually and improves the stability rate of the network.

Table 1 Route Response

No of IoT nodes	RR	PSO	SA-ACO
250	89	78	75
500	90	82	77
750	93	83	79
1000	96	89	80
1250	98	93	81
1500	100	95	85
1750	108	98	87
2000	110	99	90

Table 2 Average node connectivity time

No of IoT nodes	RR	PSO	SA-ACO
250	92	89	81
500	98	90	82
750	105	92	83
1000	109	95	89
1250	115	100	93
1500	120	100	95
1750	128	105	98
2000	130	107	99

Table 3 Degree of route imbalance

No of IoT nodes	RR	PSO	SA-ACO
250	0.25	0.23	0.22
500	0.2	0.15	0.07
750	0.1	0.1	0.05
1000	0.09	0.1	0.05
1250	0.1	0.13	0.06
1500	0.5	0.15	0.08
1750	0.8	0.18	0.09
2000	0.18	0.25	0.2

5 Conclusion

This work concentrates on constructing a multi-objective model among the service and the multiple route request generated from the service provisioning environment. This work makes use of optimal request generation and imbalance encountered by the routing nodes, as the anticipated self-adaptive ant colony optimizer model utilizes the adaptive nature to deal with the routing issues over the heterogeneous environment.

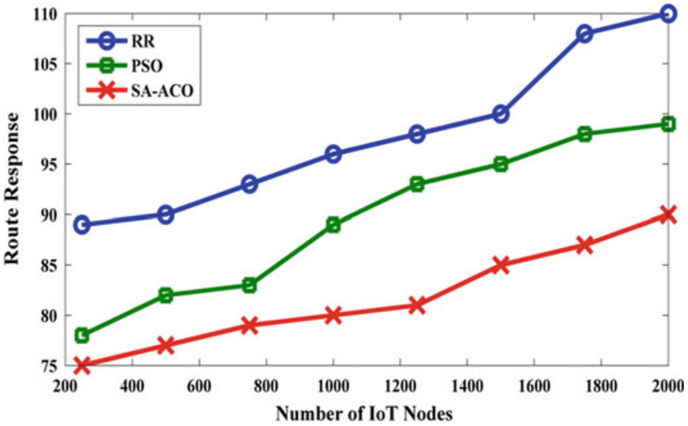


Fig. 3 Route response computation

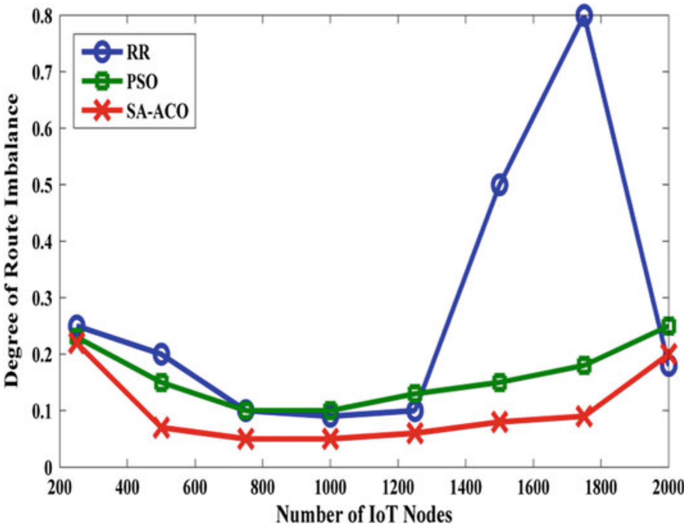


Fig. 4 Degree of route imbalance

This model is provided with a technique to deal with the superior population to identify the source and to map the route for the followers. Therefore, a stronger routing model is attained. The nodes are clustered to be adopted over the diverse network environment and assist the population to find the source in a resourceful manner. It is also provided to assist in route searching and to realize its characteristics. The simulation outcomes depict that the anticipated model can acquire the best, stronger ability and gives superior performance. In the future, hybrid optimization is applied to validate the routing performances

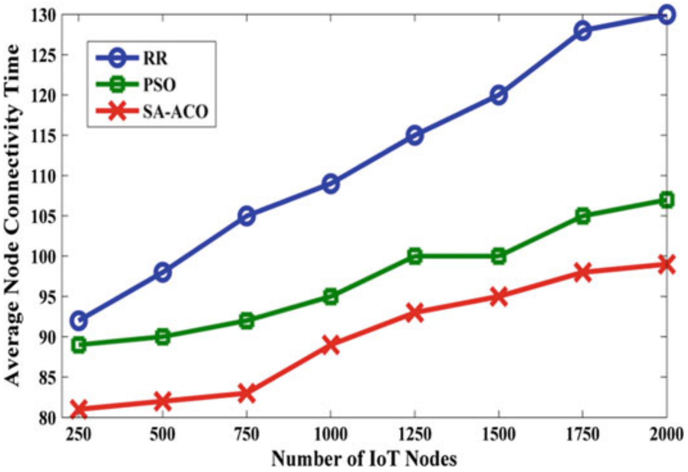


Fig. 5 Average node connectivity time

References

1. Branke J, Nguyen S, Pickardt CW, Zhang MJ (2016) Automated design of production scheduling heuristics: a review. IEEE Trans Evol Comput 20(1):110–124
2. Angelakis V, Avgouleas I, Pappas N, Fitzgerald E, Yuan D (2016) Allocation of heterogeneous resources of an IoT device to flexible services. IEEE Internet Things J 3(5):691–700
3. Chen Y, Huang JW, Lin C, Hu J (2015) A partial selection methodology for efficient QoS-aware service composition. IEEE Trans Serv Comput 8(3):384–397
4. Saravanan T, Nithya NS (2019) Modeling displacement and direction aware ad hoc on-demand distance vector routing standard for mobile ad hoc networks. Mob Netw Appl 24(6):1804–1813
5. Ding YS, Jin YL, Ren LH, Hao KR (2013) An intelligent self-organization scheme for the Internet of Things. IEEE Comput Intell Mag 8(3):41–53
6. Jin XN, Chun SJ, Jung J, Lee K-H (2017) A fast and scalable approach for IoT service selection based on a physical service model. Inf Syst Front 19(6):1357–1372
7. Saravanan T, Nithya NS (2018) Energy aware routing protocol using hybrid ANT-BEE colony optimization algorithm for cluster based routing. In: 4th IEEE ınternational conference on computing communication and automation (ICCCA), pp 1–6
8. Xu N, Ding YS, Ren LH, Hao KR (2018) Degeneration recognizing clonal selection algorithm for multimodal optimization. IEEE Trans Cybern 48(3):848–861
9. Sherubha P, Mohanasundaram N (2019) An efficient ıntrusion detection and authentication mechanism for detecting clone attack in wireless sensor networks. J Adv Res Dyn Control Syst 11(5)
10. Li L, Wang WL, Xu XL (2017) Multi-objective particle swarm optimization based on global margin ranking. Inf Sci 375:30–47
11. Wang JH, Zhang WW, Zhang J (2016) Cooperative differential evolution with multiple populations for multiobjective optimization. IEEE Trans Cybern 46(12):2848–2861
12. Sherubha P, Mohanasundaram N (2019) An efficient network threat detection and classification method using ANP-MVPS algorithm in wireless sensor networks. Int J Innov Technol Explor Eng (IJITEE) 8(11). ISSN: 2278-3075
13. Sherubha P, Mohanasundaram N, Graph based event measurement for analyzing distributed anomalies. In: Sensor networks, Sadana Academic Proceedings in Engineering Sciences, Springer (Accepted for publication)

14. Sherubha P, Mohanasundaram N, Sasirekha SP (2019) Clone attack detection using random forest and multi objective cuckoo search classification. In: IEEE ınternational conference on communication and signal processing, April 4–6, 2019, India
15. Shang RH, Dai KY, Jiao LC, Stolkin R (2016) Improved memetic algorithm based on route distance grouping for multi-objective large scale capacitated arc routing problems. IEEE Trans Cybern 46(4):1000–1013
16. Yao GS, Ding YS, Jin Y, Hao KR (2017) Endocrine-based coevolutionary multi-swarm for multi-objective workflow scheduling in a cloud system. Soft Comput 21(15):4309–4322

Optimizing Node Coverage and Lifespan of Wireless Body Area Network Using Hybrid Particle Swarm Optimization

S. Selvaraj and R. Rathipriya

Abstract Wireless body area network (WBAN) is one of the emerging wireless sensor networks for medical applications and treatments with the constrained power of numerous tiny sensors. Nowadays, many of these medical application researches focus on the low-power propagation sensing units in the mesh of the health monitoring system. This article expounds on the applicability of a discrete version of a popular benchmark swarm intelligence algorithm PSO called discrete particle swarm optimization (DPSO) and its hybrid version for energy-optimized WBAN using node coverage. DPSO-based WBAN model optimizes the node coverage for uninterrupted connectivity over the longest possible network life. The network simulator NS-2, for creating WBAN nodes, installation, connection, data transmission in the network environment is used. The simulation results showed that the proposed WBAN model has better performance in the terms of node coverage and network lifespan.

Keywords Body area sensor · Wireless sensor network (WSN) · Discrete particle swarm optimization (DPSO) · Genetic algorithm (GA)

1 Introduction

Recently, the WSN adopts any kind of ground-environment, likewise the sensor unit information with the real-time data to be used for the coverage, deployment, propagation [1, 2]. These sensor units, based on the applications it may have some prior knowledge. Generally, it consists of low-cost and low-power sensing units in small size and able to works broad range, and these applications are used as military, ground, monitoring, health devices, etc. [1–4].

Swarm intelligence is the meta-heuristic method and it gives an optimum solution to result in significant improvement in WBAN by optimizing network energy, coverage, and the level of lifespan increments [5, 6]. Figure 1 shows the WBAN structure using medical sensor data propagation. Table 1 describes the details like bandwidth and data rates of medical frequency bands.

S. Selvaraj · R. Rathipriya (✉)
Department of Computer Science, Periyar University, Salem, Tamil Nadu, India
e-mail: rathipriyar@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_61

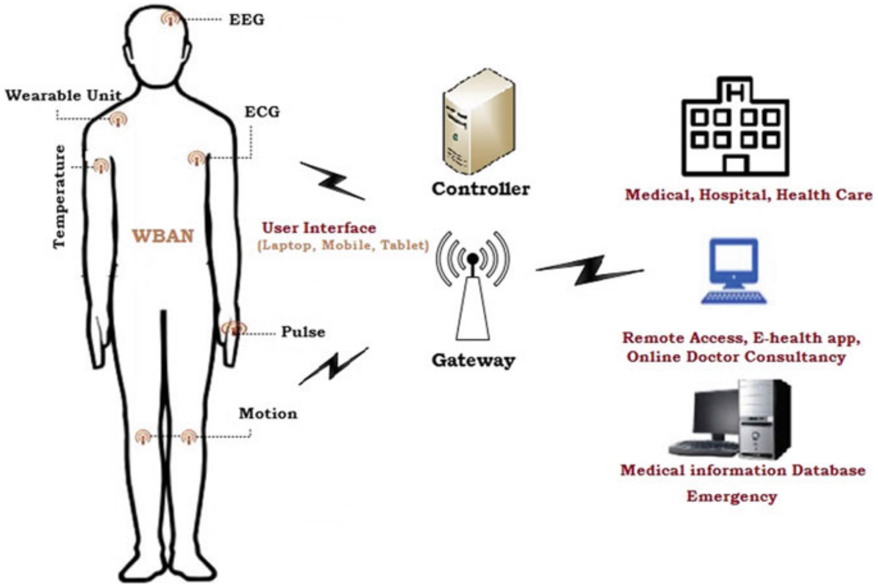

EEG
Wearable Unit
ECG
Temperature
WBAN
User Interface
(Laptop, Mobile, Tablet)
Pulse
Motion
Controller
Gateway
Medical, Hospital, Health Care
Remote Access, E-health app,
Online Doctor Consultancy
Medical information Database
Emergency

Fig. 1 Wireless body area network structure

Table 1 Worldwide medical frequency bands using bandwidth and data rates

Frequency band	Class	Bandwidth (kHz)	Channel	Data rate (kbps)
398–406	MICS (WW)	300	10	57–455
418–448	WMTS (Japan)	320	12	56–187
862–869	WMTS (Europe)	400	14	75–607

MICS Medical Implant communication service
WMTS Wireless medical telemetry service
ISM Industrial scientific medical

The full article is organized in the following manner. The introduction of WBAN is discussed in Sect. 1 and the review literature of WBAN, multi-target optimization, and swarm intelligence are explained in Sect. 2. Section 3 describes the methods and materials needed for the research work. The experimental result of the proposed work is elaborated in Sect. 4. Section 5 concludes the article with further enhancement.

2 Review Literature

This section provides the review literature needed for the research work. In [1–4], standard PSO formed as a baseline optimization method for performance testing of improvements to the technique as well as to represent PSO to a broader optimization

Table 2 Environment parameters for WBAN

Wireless technology	Illness	Parameters
WSN, AUR, Bluetooth	General illness	Cardiac output, pulse, the position of patient, meter, temperature, blood pressure, pulse-oximeter, breathing, oxygen consumption, galvanic skin reaction, ECG, movement and breath, proximity to other patients, blood glucose [1–5]
RFID, FOG Computing, GPS, Wi-Fi, WBAN	ICU patient monitoring, Dengue, chronic illness	ECG, oxygen levels, temperature, blood pressure, pulse [1–5]

community. In the second stage, the best classifier variants improved in terms of attribute type support and temporal complexity. These works defined and addressed two data-related issues that could affect the efficiency of particle swarm optimization: high-dimensional data sets, mixed attribute data, and proposed solution to each of these issues including recent improvements by a PSO algorithm, with the latest developments that helped to improve performance on standard measures to extend the first particle swarm optimization [5]. Table 2 showed the parameters used and the type of illness of the most commonly used wireless technology [9].

The solution transfers from one population to another through certain assessment procedures in the population-based method. GA uses a genetic system such as selection, mutation, and crossover, while for the assessment process, PSO uses swarm behavior such as updating location and speed [6].

3 Methods and Materials

This section describes the required methods and materials for the proposed DPSO and its hybridization for optimizing the node coverage and network lifespan. A combination of optimization algorithms is used in the many works to optimize the parameters for the numerous problems of various domains.

3.1 Genetic Algorithm

The genetic algorithm creates a paradigm for evolution simulating genetically determined Darwinians and the normal cycle of elimination. The transition procedure ensures that the entire population moves to the global optimum through better chromosomes. Mutation operations maintain the diversity of the population and prevent

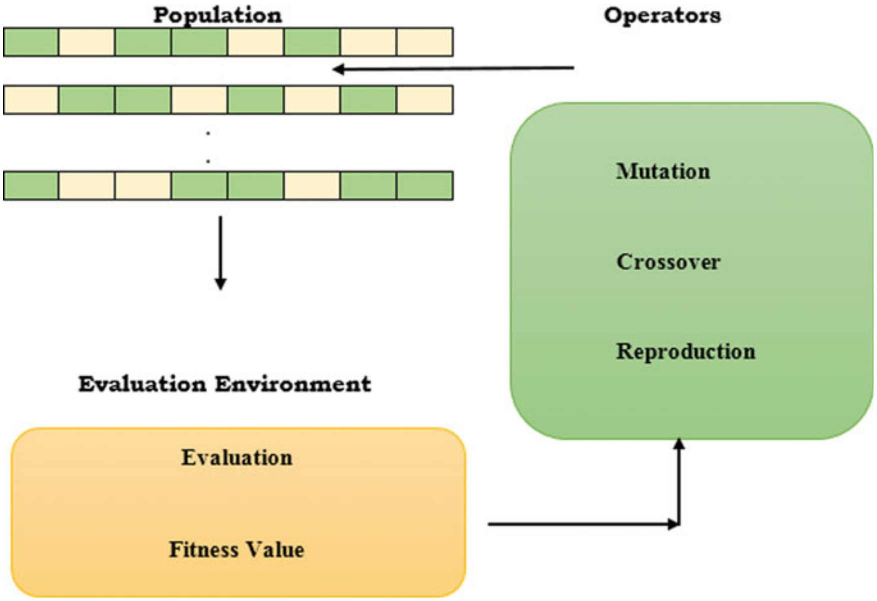


Fig. 2 Flow of GA

the population from dropping in the optimal location. Figure 2 shows the flow of GA [6, 7].

3.2 Discrete PSO

To replace the velocity concept, the study presented and modified several operators based on the swap operator and the swap series concept. It is known to be a special DPSO to solve discrete problems. DPSO begins the process of searching using a randomized group of particles and the flow presents in Fig. 3. Then, particles are used to defined pbest and gbest [8].

The critical difference of DPSO is that the particulate velocity and position. The variation of probabilities characterizes these, and the particles generated by an integer in [0, 1]. A particle flies, therefore, in a search space confined to zero and one. The velocity is limited, and the interval has [0, 1].

The sigmoid function transformation as $S\ (v_i\ (t+1))$ is shown in the Eq. 1 can be used to limit the speed.

$$S_{sig}(V_i(t+1)) = \frac{1}{1+e^{(V_i(t+1))}} \tag{1}$$

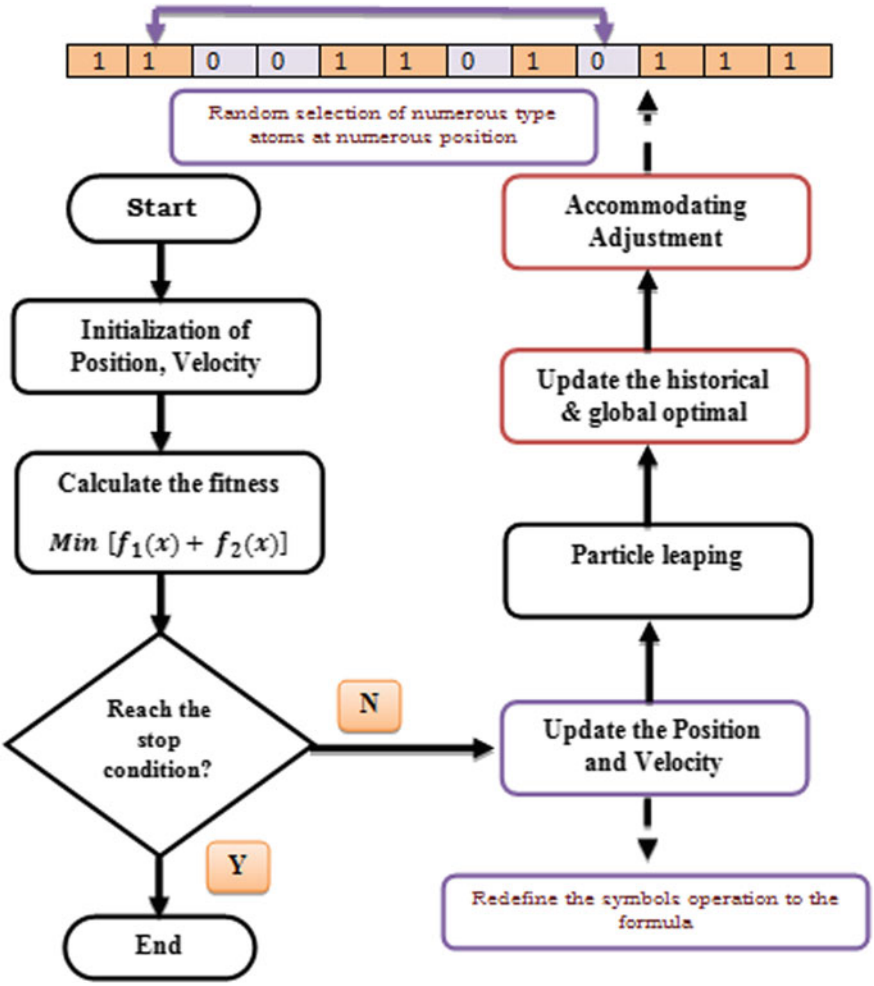

1 1 0 0 1 1 0 1 0 1 1 1
Random selection of numerous type atoms at numerous position
Start
Initialization of Position, Velocity
Calculate the fitness
$Min\ [f_1(x) + f_2(x)]$
Reach the stop condition?
N
Y
End
Accommodating Adjustment
Update the historical & global optimal
Particle leaping
Update the Position and Velocity
Redefine the symbols operation to the formula

Fig. 3 Flow of DPSO

The updated position is determined using Eq. 2, where r_3 is assigned to a uniform random value and the range is [0, 1].

$$x_{iD} = \begin{cases} 1 \text{ if } r_3 < (v_i(t + 1)) \\ 0 \qquad \text{otherwise} \end{cases} \tag{2}$$

3.3 Simulated Annealing

An analogy to annealing of ideal crystals in thermodynamics is based on a simulation annealing (SA). The algorithm is designed to maximize the molecules' thermal movement at fixed temperature simulations. Thus, for instance, the term 'temperature' used to name a crucial control parameter is derived from thermodynamical SA. In SA operations, the temperature parameter is high continuously. Rapid cooling creates defects that do not achieve sufficient energy levels in the crystal structure [8–11]. The mechanism imitates this process while optimizing the parameters in SA. The possibility of such approval is dependent on temperature and should be negligible at low temperatures. A significant function of SA's approach is an opportunity to consider a momentarily bad option [12–15]. Figure 4 illustrates the complete workflow of the SA algorithm for better understanding.

4 HPSO-Based Energy Model for Extended WBAN Lifespan

The major drawback of the DPSO is that in high-dimensional space, it is easy to collapse into local optimum and has a poor convergence rate in the iterative method. To overcome the above-said problem, DPSO is hybridized with SA to develop a hybrid DPSO algorithm (namely HPSO) to avoid local optima. The HPSO-based energy efficiency model is proposed for WBAN. It is a population-based optimization method. Algorithm 1 describes the HPSO-based energy model for extended WBAN lifespan [13–15].

The WBAN used in this study has three different layers namely, the sensing layer split into an outdoor ground sensor, indoor/room sensor, and inside-wearable WBAN sensors [16–20]. Network Layer has the controller to sense and control all sensing units like a gateway of WBAN. This layer is connected to public, private, industry clouds, and storage. Finally, it has the interface layer responsible for data aggregation, visualizations, and the analytical process of body sensors.

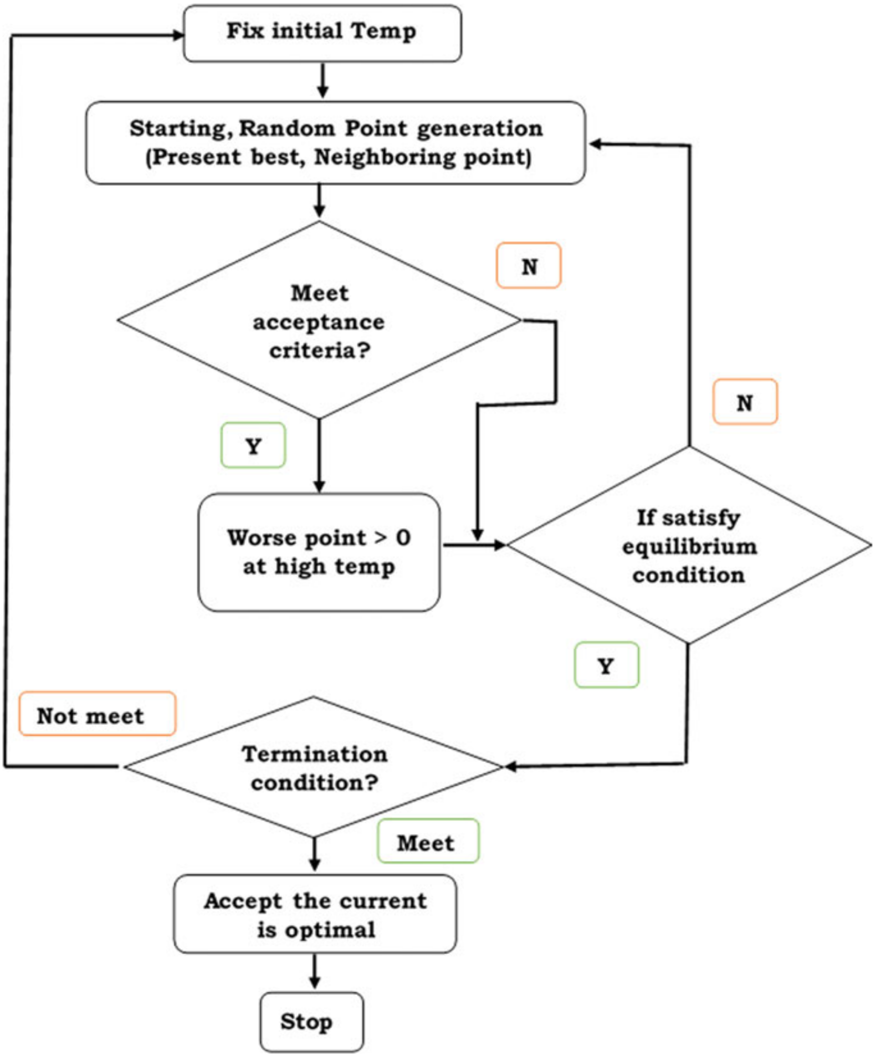

Fix initial Temp
Starting, Random Point generation (Present best, Neighboring point)
Meet acceptance criteria?
N
Y
Worse point > 0 at high temp
N
If satisfy equilibrium condition
Y
Not meet
Termination condition?
Meet
Accept the current is optimal
Stop

Fig. 4 Flow of SA

Algorithm 1: HPSO Based Energy Model for Extended WBAN Lifespan

```
Input: Set of initial active nodes, Energy status of the nodes in WBAN
Output: best, the globally optimal set of active nodes

Step 1. Set initial particles                    // Initialization
Step 2. Initial the parameters (v  and p) of each particle
Step 3. Set pbest and gbest
Step 4. f=evaluate()
Step 5. Repeat
            for each candidate
               SA(best(particle)
               update position, pbest and gbest and
            f1=evaluation()
             if  f1 > f
            Update position, pbest and gbest
            end if
            End for
        SA(best(particle)
        Update best
        Update v
        Update p
        until facing the stopping criteria
Step 6. Return best as optimal global particle
```

The main objective of this proposed model is to identify the optimal global subset of an active node for communication in WBAN with minimal network operational energy for an extended network lifespan. HPSO is initialized with an initially random set of active nodes in the network. HPSO-WBAN energy model is used to increase the network lifespan of WBAN by managing the nodes' remaining energy in the network and the total number of nodes alive during the simulation for successful network communications.

4.1 Fitness Function

A multi-objective maximization fitness function evaluated in the HPSO hybrid algorithm optimizes the life cycle of the WBAN using an active set of nodes. It helps to ensure that the condition envisaged with each response is acceptable. To desire, network coverage is a maximum probability of points in the target region that can be sensitive and well-defined by Eq. 3. Table 3 illustrates the design goals and criteria taken for the study. Table 4 shows the network simulation setup of the proposed work.

$$\max f_A(x) = \frac{A_{TS}}{A}. \tag{3}$$

Table 3 Proposed design perception

Design goals	Criteria	Challenges
Propagation (sense to network)	Zig-bee, Wi-Fi, Bluetooth	Diversity, expense, protection, privacy, data acquisition, and truthfulness
Propagation (network to cloud)	CoAP/HTTP/FTP	
Energy source	Battery powered	
Latency	50 s for data	
Storage	Cloud (public, private, industry)	

Table 4 Simulation parameters

Parameters	Value
Area	500 × 500
No. of relay nodes	8
No. of sensor nodes	100
Transmission of relay nodes	50 m
Propagation	Two Ray Ground
Network interface type	Wireless Phy
Traffic type	CBR, FTP
IEEE 802.15.4 standard	Default values
Simulation time	1000 s
Initial energy	50 J
Energy threshold	20% of the initial energy
Sensors	EEG, ECG, pulse, motion, blood pressure, blood glucose

A_{TS} is the active sensors in the target region, and A is the region.

5 Results and Discussions

This study is to analyze the efficiency of the GA, DPSO, and hybrid for node coverage optimization. The estimation time is seen in Figs. 5, 6, and 7 for DPSO, GA, hybrid.

5.1 Packet Delivery Ratio

Figure 6 shows the packet delivery, and the *x*-axis indicates the network node level; the *y*-axis is the delivery ratio of the particular nodes in the region. The packet

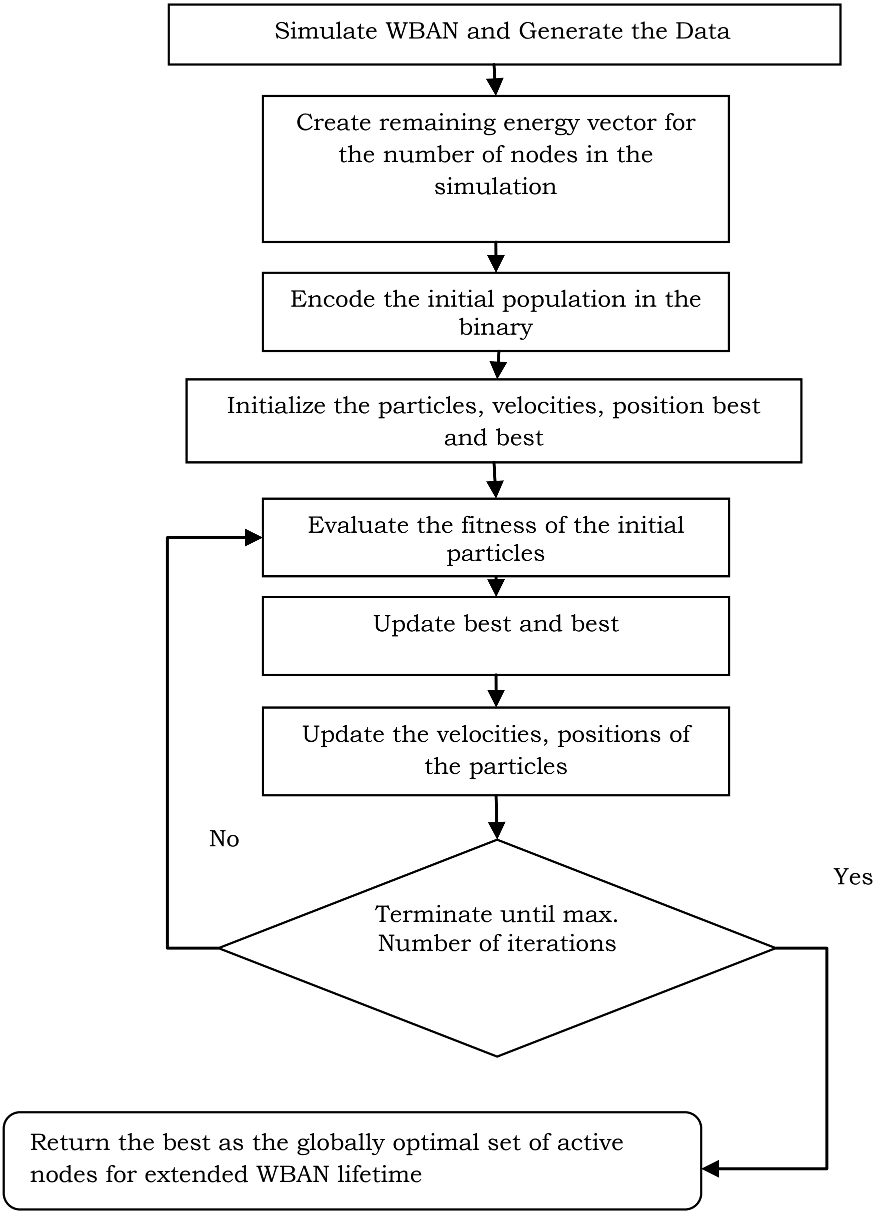

Simulate WBAN and Generate the Data
Create remaining energy vector for the number of nodes in the simulation
Encode the initial population in the binary
Initialize the particles, velocities, position best and best
Evaluate the fitness of the initial particles
Update best and best
Update the velocities, positions of the particles
No
Yes
Terminate until max. Number of iterations
Return the best as the globally optimal set of active nodes for extended WBAN lifetime

Fig. 5 A workflow of HPSO-based energy model for extended WBAN lifespan

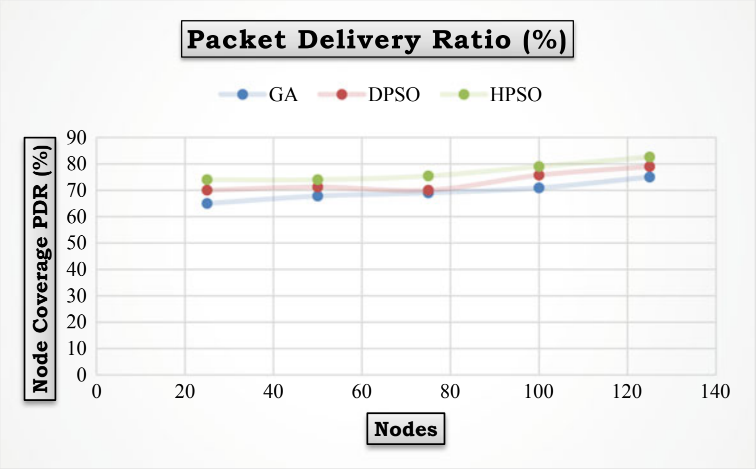


Fig. 6 Packet delivery ratio

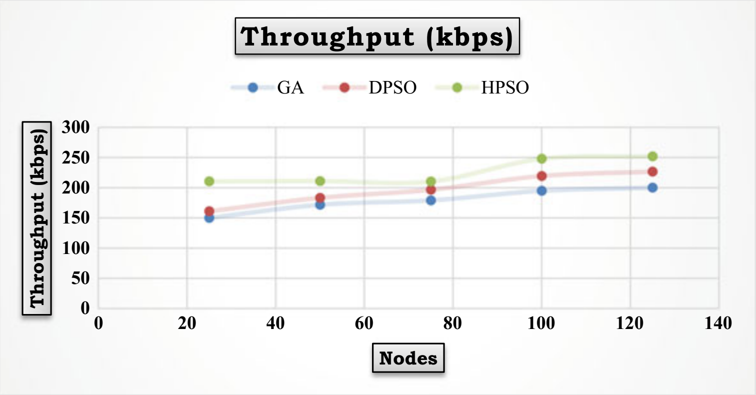


Fig. 7 Throughput

delivery ratio is 17, 24, and 31% increases in GA, DPSO, and HPSO. It needs a cumulative number of packages sent and some packets received to determine the packet distribution ratio. The packet rating is sustained at 90 percent, even though network size exceeds 100 nodes, with higher efficiency compared to GA, DPSO, and HPSO. Figure 6 represents the packet delivery ratio with a percentage.

$$\textbf{Packetdeliveryratio} = \frac{\sum \text{PacketsReceived}}{\sum \text{PacketSent}}$$

5.2 *Throughput*

Figure 7 shows the throughput level of WBAN and results showed DPSO, hybrid PSO ratio of the throughput is high compare with the GA, DPSO. It is comprehensible from Fig. 7, the throughput value for 150 nodes in the simulated WBAN is 08 s. Similarly, the throughput value for 120 nodes in the simulated WBAN is 52 s. Therefore, the throughput value is openly relative to the network range.

5.3 *End-To-End Delay*

Figure 8 represents the end-to-end delay performance of GA, DPSO, and hybrid PSO for optimal node coverage in WBAN. *X*-axis represents the number of nodes, whereas the *y*-axis represents the end-to-end delay time in second. Hybrid BPSO-based node coverage algorithm has lower end-to-end delay value when compared to GA and BPSO similarly, the EED value for 100 nodes in the simulated WBAN is 57 s. Therefore, the end-to-end delay value is directly proportional to the network size.

$$\text{End - to - End Delay} = \text{Received time} - \text{Sent time}$$

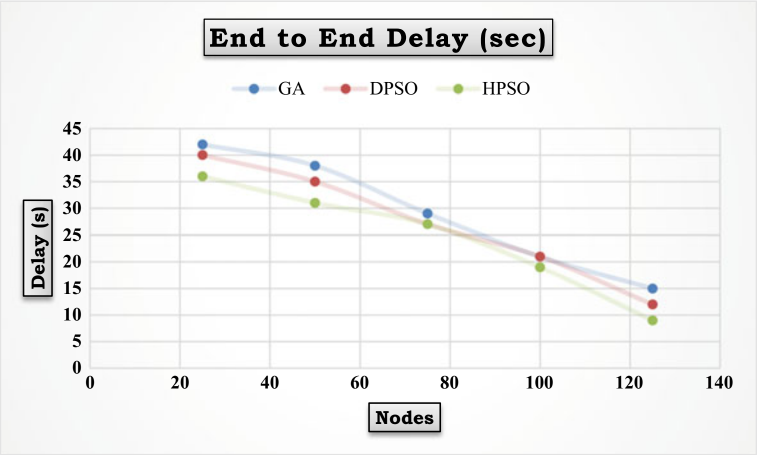


Fig. 8 End-to-end delay

6 Conclusion and Future Scope

In this work, the coverage optimization of nodes in the WBAN is accessed through QoS parameters such as packet transfer rate, end-to-end delay, and performance. The HPSO algorithm has eliminated the local convergence impact of HPSO and GA algorithms and increases the consistency between exchanging and filtering information to find the best solution in the search field. The performance of the HPSO and standard DPSO algorithm has been studied in the simulation environment for identifying the optimal node coverage nodes for seamless communications. From the simulation results, it can be seen that the HPSO-based node coverage algorithm for WBAN performs better than the DPSO. In the future, different hybrid mechanisms will be identified, proposed, and studied for the same environment.

References

1. Sarkar A, Maunder S (2015) Path loss estimation for a wireless sensor network for application in ship. Int J Comput Sci Mob Comput
2. Astudy, Haryono T (2016) Novel binary PSO algorithm based optimization of transmission expansion planning considering power losses. In: International conference on innovation in engineering and vocational education, IOP Publishing, IOP Conf. Series: Materials Science and Engineering, vol 128, p 012023. doi:https://doi.org/10.1088/1757-899X/128/1/012023
3. Dhamodharavadhani S (2015) A survey on clustering based routing protocols in Mobile ad hoc networks. In: 2015 international conference on soft-computing and networks security (ICSNS). doi: https://doi.org/10.1109/icsns.2015.7292426
4. Sivabalan S, Dhamodharavadhani S, Rathipriya R (2019) Opportunistic Forward routing using bee colony optimization. Int J Comput Sci Eng 7(5):1820–1827. https://doi.org/10.26438/ijcse/v7i5.18201827
5. Nguyen BH, Xue B, Andreae P, Zhang M, A new binary particle swarm optimization approach: momentum and dynamic balance between exploration and exploitation. IEEE Trans Cybern
6. Jamil F, Iqbal MA (2019) Adaptive thermal-aware routing protocol for wireless body area network. In: International conference on broadband and wireless computing, communication, and applications
7. Kaur HP (2015) Cost-based efficient routing for wireless body area networks. Int J Comput Sci Mob Comput
8. Dhamodharavadhani S, Rathipriya R (2020) Enhanced logistic regression (ELR) model for big data. In Garcia Marquez FP (ed) Handbook of Research on Big Data Clustering and Machine Learning, pp 152–176, IGI Global. http://doi.org/10.4018/978-1-7998-0106-1.ch008
9. Mohd Kaleem M (2014) Energy consumption using network stability and multi-hop protocol for link efficiency in wireless body area networks. J Comput Eng 16 113:120
10. Mohamad MS, Omatu S, Deris S, Yoshioka M, Abdullah A, Ibrahim Z (2013) An enhancement of binary particle swarm optimization for gene selection in classifying cancer classes. Algor Molecul Biol 8:1510
11. Romesh Singh M (2019) Development of efficient multi-hop protocols for wireless body area network. Int J Innov Technol Explor Eng
12. Sivabalan S, Dhamodharavadhani S, Rathipriya R (2020) Arbitrary walk with minimum length based route identification scheme in graph structure for opportunistic wireless sensor network. Swarm Intell Resour Manage Internet Things 2020:47–63. https://doi.org/10.1016/b978-0-12-818287-1.00006-1

13. Sivabalan S, Rathipriya R (2017). Slot scheduling Mac using energy efficiency in ad hoc wireless networks. In: 2017 international conference on inventive communication and computational technologies (ICICCT). doi: https://doi.org/10.1109/icicct.2017.7975234
14. Thangavel K, Bagyamani J, Rathipriya R (2012) Novel hybrid PSO-SA model for biclustering of expression data. Proc Eng 30:1048–1055. https://doi.org/10.1016/j.proeng.2012.01.962
15. Rathipriya R, Thangavel K (2012) A discrete artificial bees colony inspired biclustering algorithm. Int J Swarm Intel Res 3(1):30–42. https://doi.org/10.4018/jsir.2012010102
16. Dhamodharavadhani S, Rathipriya R (2020) Variable selection method for regression models using computational intelligence techniques. In Ganapathi P, Shanmugapriya D (ed) Handbook of research on machine and deep learning applications for cyber security (pp. 416–436). IGI Global. https://doi.org/10.4018/978-1-5225-9611-0%2Ech019
17. Dhamodharavadhani S, Rathipriya R (2021) Novel COVID-19 mortality rate prediction (MRP) model for India using regression model with optimized hyperparameter. J Case Inform Technol (JCIT), 23(4):1–12. https://doi.org/10.4018/JCIT.20211001.oa1
18. Selvaraj S, Rathipriya R (2019) Energy efficiency in wireless body area networks using path loss model. Int J Comput Sci Eng 7(5):1695–1700. https://doi.org/10.26438/ijcse/v7i5.16951700
19. Dhamodharavadhani S, Rathipriya R (2018) Region-wise rainfall prediction using mapreduce-based exponential smoothing techniques. In: Advances in intelligent systems and computing advances in big data and cloud computing. 229–239. https://doi.org/10.1007/978-981-13-1882-5_21
20. Dhamodharavadhani S, Rathipriya R, Chatterjee JM (2020). COVID-19 mortality rate prediction for India using statistical neural network models. Frontiers Public Health 8. https://doi.org/0.3389/fpubh.2020.00441

Collective Examinations of Documents on COVID-19 Peril Factors Through NLP

E. Laxmi Lydia, Jose Moses Gummadi, Chinmaya Ranjan Pattanaik, B. Prasad, CH. Usha Kumari, and Ravuri Daniel

Abstract The outbreak of the novel COVID-19 virus is identified across all experimental scientific tests that assist victims to fight against the pandemic situation. The problem seems to have a large number of scientific COVID-19 articles with different risk factors. The quick identification of documents allows the processing and interpretation of inevitable essential knowledge for investigators. This article provides a solution by creating an unsupervised framework for the interpretation of clinical trials over COVID-19 risk factors with a diverse range of articles related to vaccines and treatments from a large corpus of documents. It also provides practical informative knowledge regarding COVID-19 risk factors and helps researchers to enable any single author to obtain appropriate information. The present application uses artificial intelligence, natural language processing approaches, incorporated throughout

---

E. Laxmi Lydia (✉)
Department of Computer Science and Engineering, Vignan's Institute of Information Technology (A), Visakhapatnam, Andhra Pradesh, India
e-mail: elaxmi2002@yahoo.com

J. Moses Gummadi
Department of CSE, VFSTR (Deemed to be University), Guntur, India
e-mail: josemoses@gmail.com

C. Ranjan Pattanaik
Department of Computer Science and Engineering, Ajay Binay Institute of Technology, Cuttack, Odisha, India
e-mail: chinmaya.pattnaik@rediffmail.com

B. Prasad
Department of Information Technology, Vignan's Institute of Information Technology (A), Visakhapatnam, Andhra Pradesh, India
e-mail: arjunprasad.bode@gmail.com

CH. Usha Kumari
Department of ECE, GRIET, Gokaraju Rangaraju Institute of Engineering & Technology, Hyderabad, India

R. Daniel
Department of Computer Science and Engineering, Bapatla Engineering College (Autonomous), Bapatla, India
e-mail: danielravuri@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_62

the search engines, to search for keywords to classify categories with normalized linguistic data. The text data are instead parsed in phrases and thresholds the text with recognition of data frame components with relevant outcomes.

Keywords SARS-CoV-2 · COVID-19 · MERS-CoV · Cytotoxic chemotherapy · Natural language processing (NLP) · Artificial intelligence (AI)

1 Introduction

The newly updated incidence of coronavirus (2019-nCoV) in Wuhan has become a worldwide threat to humans. The world within 2 months after its arrival in December 2019, nearly 70,000 people was infected with such a newly emerging coronavirus that resulted in around 1700 deaths. In response to the increasing number of patients infected daily, this new virus is estimated to be a source of a major pandemic. Two such big coronavirus epidemics have previously been encountered in the globe. In 2002–2003, a new coronavirus named SARS broke out and spread into other countries, like Vietnam, Canada, mostly in China and Hong Kong. About 8000 cases have been filed initially with overall 774 deaths worldwide [1]. From then, SARS death cases were reduced. Nevertheless, among several bat species with the possibility to infect human cells, a new SARS-like coronavirus was found [2]. Throughout the Middle East Asian countries, after 10 years a new variety of coronavirus known as MERS was identified. A total of 1308 laboratory-confirmed cases, along with 449 individuals treated as female, were registered in Saudi Arabia between November 2012 and February 2016 [3]. 400 of the 449 female patients; 179 were reproductive (15–45 years) age details were available and 16 (i.e., 8.9 percent) of the 179 cases. Including its 1308 MERS-CoV confirmed cases, five MoH cases of pregnancy have been identified [4].

Such studies revealed that the Indian economy at a low rate, coronavirus infections occur more in India. One of the researcher Suresha revealed that out of 9 virus-infected cases occurred among the 1706 occurrences (1,04%), 4 HCoV-OC43 cases, 3 cases HCoV-NL63 were positive, and 2 HCoV229E were positive [5]. According to the AIIMS study, New Delhi demonstrated that 17.8% of infants suffer from coronavirus infection. Most of these clinical trials produced certain survey factors which might imply the potential propagation route or risk factors. A broad range of environmental viruses including the coronavirus, hantaviruses, lyssaviruses, Lassa virus, Rabies, Nipah, Henipavirus, Ebola virus, SARS coronavirus, and Marburg virus have been reported from bats [6]. Bats are also considered a natural host. SARS and MERS have been identified to be from a bat in early coronavirus epidemic situation. Nationally and internationally, there are over 1200 bat species of varying sizes, ranges, and ecosystems. In the Indian subcontinent, almost 128 bat species were described. The most unusual, formed and rendered bat (Kerivoula picta), rare bat of Salim Alis (Latidens salimalii); the Indian flying fox (Pteropus giganteus) was identified as unusual bats found in the world. Such bats move from higher Himalayan areas, northwestern deserts to tropical forests in the eastern region. Indian flying fox is the dominant species in India that have twelve types of flying foxes. Throughout the world, only three are popular the flying Indian fox, Rousettus leschenaultia, and Cynopterus sphinx. Certain habitats are uncommon and can only be found on peaks and islands of Andaman and Nicobar in the Thailand region [7].

The population of India is at greater or less risk with an outburst of coronavirus and got alerted to prevent the death cases due to coronavirus pandemic. Drastic action is expected to grow substantially to prevent any global epidemic of coronavirus. To establish the diseases of coronavirus spreading throughout this population, a new strain needs to be governed. In underlying mechanisms of virus transmission, research needs to be carried out. Providing essential insights into the sequence of transmission, multiple strains of coronavirus have different receptors and receptor biology. Strategies to recognize receptors for unidentified viruses and medicines intended to stimulate the interconnection between virus receptors. Academic research on the human biology and disease progression of coronavirus, the new drugs too should be pursued. COVID-19 was initiated in the development and manufacture of vaccines. GSK (Vaccine giant) and Jenner Institute, in Great Britain, intended to provide the same framework for the preparation of MERS-CoV flu shots. India is indeed one of the leading manufacturers of medicines, and several provinces use those manufactured vaccines. India often has a solid system for the short-term development of large-scale vaccinations. Resources that manufacture medicines use the formulation of the coronavirus which includes initial research papers as a preparation to prevent an unexpected occurrence to assess the efficacy and reliability of the vaccine in India.

2 Literature Survey

India is experiencing a rapid epidemiological transformation with an emerging occurrence of chronic sicknesses that has a population of nearly one billion with lower-middle-income countries (i.e., 68, 85 percent). In 2012 [8], cardiac disease and stroke have been among the top three major causes that contributed to the decline in premature global deaths, and hypertension, another of the strongest cardiovascular risk factors, gives rise to 45% of heart diseases [9]. In 1990 India, chronic diseases reported as 3.78 million (40.4%), and by 2020, this has crossed 7.63 million (66.7%). Hypertension has become a major health issue in most Indian regions. 10.8% of deaths and 4.6% of disabilities are estimated in India [10]. The higher efficacy of hypertension in India was related to measures like aging, current drinking, stress, and anxiety level and body mass of more than 90 cm in coastal Pondicherry [11].

Bartwal et al. [12] noticed that hypertension to be 41.7%, and hypertension is linked to age development, the household background of hypertension, increased salt consumption, a mixed diet, a rise in the thickness of the neck, body mass index, and waist–hip ratio. Interestingly, Kokiwar et al. [13] identified that alcohol consumption amongst rural communities in Central India hasn’t been tied to hypertension while factors as the upper social class, sedentary exercise habits; misuse of tobacco, and diabetes were significantly correlated in hypertension [14].

The framework of World Health Organization (WHO) has been used [15] for the assessment of behavioral risk factors like drinking, eating habits, physical activity, smoking, and obesity. A weight of kilograms divided by height (kg/m^2) is derived from body mass index calculation. To represent the involvement of behavior classified by each risk factor was labeled as 0/1. A risk factor index has reportedly been managed to create by summarizing the factors. The spike glycoprotein or antibody with the foremost growth factor of viral infection is aimed at particularly strategies to fix Cov vaccines and medicines. Few persons demonstrated adequacy in vitro studies and expanded to parallelize animal or human studies, with insufficient use to combat COVID-19 infection [16].

Diagnosing COVID-19, coronavirus-caused disease (sars-cov-2) with several vaccines (MADs/RADs) has been interpreted scientifically. They have analyzed clinical results of how to use antiretroviral treatment for coronavirus detection and care [17]. The clinical results from SARS, MERS, COVID-19 patients trained to treat antiretroviral have been screened in primary studies. Two randomized clinical trials, 24 observational studies were administered from a previous sample of 433 names, including clinical evidence antiretroviral. Including its 21 laboratory experiments demonstrating patient outcomes, three trials have always been conducted on SARS patients, six were conducted on MERS patients, and 12 were performed on COVID-19 patients. 3 out of 361 patients undergoing lpv/r died in randomized trials; scientific proof remained poor. The beneficial role of lpv/r has been defined as post-exposure prophylaxis by three studies [18].

Articulating the current preclinical and clinical data from initial studies emphasize several significant modified features of SAR-COV-2 that further distinguish between SAR-COV and middle eastern coronavirus syndrome (MER-COV), such as large variability of infection assessment. Recent clinical trials have shown that a variety of such drugs are effective, such as favipiravir, an antiviral within that vast array that influences viral replication, and hydroxychloroquine, a converted antimalarial drug that impairs with the endosomal access system of the virus. They have anticipated that the global pandemic disaster will lead to more systematic pharmaceutical techniques to design provides a broad data analysis [19].

For the control of this immediate and lethal disease, no unique anti-virus medicines or vaccines were generated. Modern screening tests are also carried with underlying mechanisms, the therapeutic effect of TCM, and the reorganization of new natural anticoronaviral compounds [20].

The probability of developing extreme and even deadly respiratory diseases is relatively high for cancer and transplant patients with COVID-19, particularly as they may be treated with immune or immune-stimulant drugs. This investigation focuses on the impact of all these medicines on host immunity to COVID-19. Various approaches for immune-suppressing or -stimulating drugs use Ovid Medline. They are cytotoxic chemotherapy, low-dose steroids, tnfα, il-6, jak inhibitors, il-1 blockade, mycophenolate, tacrolimus, anti-cd20, and ctla4-ig. For performance, 89 research studies have been conducted [21]. Many epidemiological studies have proven that the efficacy of hypertension elevated among COVID-19 patients affected by COVID infection as a high-risk factor [22].

Based on the most frequent use of COVID-19 drugs chloroquine and hydroxychloroquine (HCQ), patients were identified to have improvement. Taking into account the low-risk factors, the extensive further association of other diseases shows cost-effectiveness and easy access in India. Researchers suggested that both prescribed drugs are worthy of an immediate clinical track trial, which can be carefully considered as observational medicines for clinical use [23].

Evidence of patient protection in the use of chloroquine and hydroxychloroquine is obligatory for many years. When experimenting with COVID-19, the main target of chloroquine and hydroxychloroquine under different viral conditions has been limited [24].

In a standard procedure, investigators may have randomized controlled monitoring precautions for the effects of chloroquine and hydroxychloroquine—medication or in conjunction with other medications—as eligibility criteria under several concurrent clinical evaluations. In case of randomized infections caused by another coronavirus, such as MERS-COV and SARS-COV, and unrandomized studies of COVID-19, randomized trials would be tested if no support for important findings appears specifically observed in randomized studies or if the proof is small or very low [25].

The only remedy to the sudden infectious disease epidemic can be repositioning of medications. A total of 21 objectives have been assessed toward structure-based virtual libraries from zinc drugs and existing dataset of natural products using selective ligand screening. A collection of 78 most often-used antiviral medicines has also been developed that effectively addresses those commercially available and drug trials for SARS-COV-2. Requirements and future drug targets for these drugs have been expected. This analysis includes different vaccines and aims for further in vitro and in vivo SARS-COV-2 treatments. The new developments in existing clinical trials of these drugs show potential drug repositioning strategies for the treatment of SARS-COV-2 infections [14].

3 Methodology

Following is the process which enables medical research group, communities, and policymakers to easily review recent studies and developments in a specified area of knowledge. Interactive search queries allow practitioners to search for COVID-19-related knowledge. The full-text scientific study is also integrated into the final output frame and can be clicked directly.

(a) Initially, a variety of abstracts were particularly been analyzed by COVID-19 and its variations. Due to the extremely huge scale of the information.

(b) SciSpacy on the abstracts used to perform tokenization and data analysis includes stemming, lemmatization, and stop-word elimination. ScispaCy is perhaps a Python package that comprises SpaCy modeling for life sciences, research-based or pharmacological text processing.

(c) The manuscripts (abstracts) have been put into a TF-IDF model for TF-IDF estimation.

(d) This identified abstract was calculated against the cosine similarities of the dynamic user application.

(e) Finally, it filters the most related documents and reveals the ten most important documents.

A. **Procedure**

Install the SpaCy pipeline with a broad vocabulary and 600 k word vectors for biomedical data.

Step 1: Import packages such as NumPy, pandas, scisspacy, spacy, phrasematcher, CountVectorizer, TQDM.

Step 2: Load the metadata in the abstract of the article designers choose.

Step 3: Collect all related research papers relevant to COVID-19 and its derivatives.

Step 4: Load sci model and initiate tokenizer (to discard stop words and lemmatizing words).

Step 5: Perform training process through TF-IDF vectorizer over research papers.

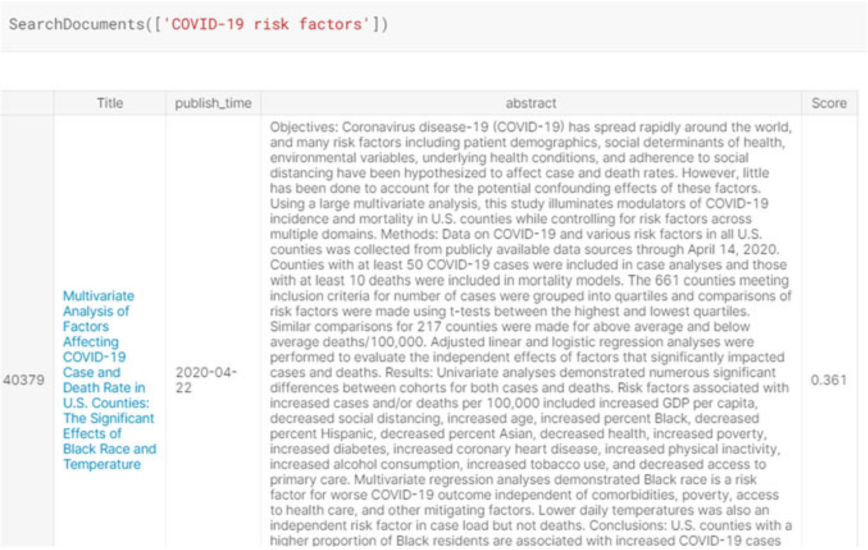
Step 6: Characterize the similarity of cosine to procure top n documents.

Step 7: Classifies method to obtain the top ten search query documents and a process to respond to every question.

Step 8: Set methods to screen the data in the table.

4 Result Analysis

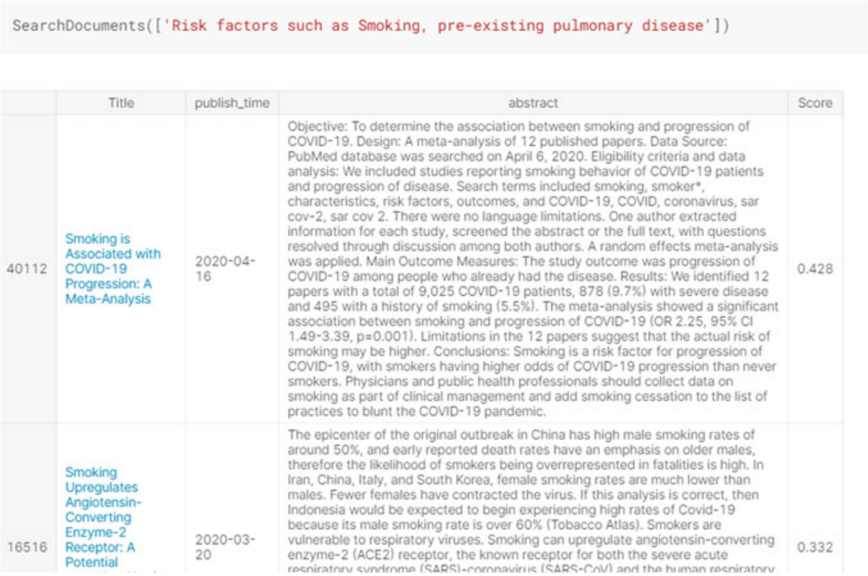
The following experimental results were carried out based on the Kaggle dataset, i.e., COVID-19 open research dataset challenge. This article provides a platform for extracting and classifying relevant information that could be enhanced in future using highly qualified interpretation in the absence of expert feedback. Data concerning risk factors, associated with COVID-19 such as smoking, additional lung disorders, bacterial infections, as well as other co-organisms, socio-cultural, and psychological factors to influence the actual impacts and similarities of the virus; preterm infants and nursing mothers to measure whether heterogeneous respiratory/viral infections create the virus quite easily transmitted are examined. Figure 1 shows the top n document abstracts related to COVID-19 risk factors. Figure 2 shows the top n document abstracts related to risk factors like smoking and pre-existing pulmonary disease; Fig. 3 shows the top n document abstracts related to COVID-19 heart risks.



```
SearchDocuments(['COVID-19 risk factors'])
```

	Title	publish_time	abstract	Score
40379	Multivariate Analysis of Factors Affecting COVID-19 Case and Death Rate in U.S. Counties: The Significant Effects of Black Race and Temperature	2020-04-22	Objectives: Coronavirus disease-19 (COVID-19) has spread rapidly around the world, and many risk factors including patient demographics, social determinants of health, environmental variables, underlying health conditions, and adherence to social distancing have been hypothesized to affect case and death rates. However, little has been done to account for the potential confounding effects of these factors. Using a large multivariate analysis, this study illuminates modulators of COVID-19 incidence and mortality in U.S. counties while controlling for risk factors across multiple domains. Methods: Data on COVID-19 and various risk factors in all U.S. counties was collected from publicly available data sources through April 14, 2020. Counties with at least 50 COVID-19 cases were included in case analyses and those with at least 10 deaths were included in mortality models. The 661 counties meeting inclusion criteria for number of cases were grouped into quartiles and comparisons of risk factors were made using t-tests between the highest and lowest quartiles. Similar comparisons for 217 counties were made for above average and below average deaths/100,000. Adjusted linear and logistic regression analyses were performed to evaluate the independent effects of factors that significantly impacted cases and deaths. Results: Univariate analyses demonstrated numerous significant differences between cohorts for both cases and deaths. Risk factors associated with increased cases and/or deaths per 100,000 included increased GDP per capita, decreased social distancing, increased age, increased percent Black, decreased percent Hispanic, decreased percent Asian, decreased health, increased poverty, increased diabetes, increased coronary heart disease, increased physical inactivity, increased alcohol consumption, increased tobacco use, and decreased access to primary care. Multivariate regression analyses demonstrated Black race is a risk factor for worse COVID-19 outcome independent of comorbidities, poverty, access to health care, and other mitigating factors. Lower daily temperatures was also an independent risk factor in case load but not deaths. Conclusions: U.S. counties with a higher proportion of Black residents are associated with increased COVID-19 cases	0.361

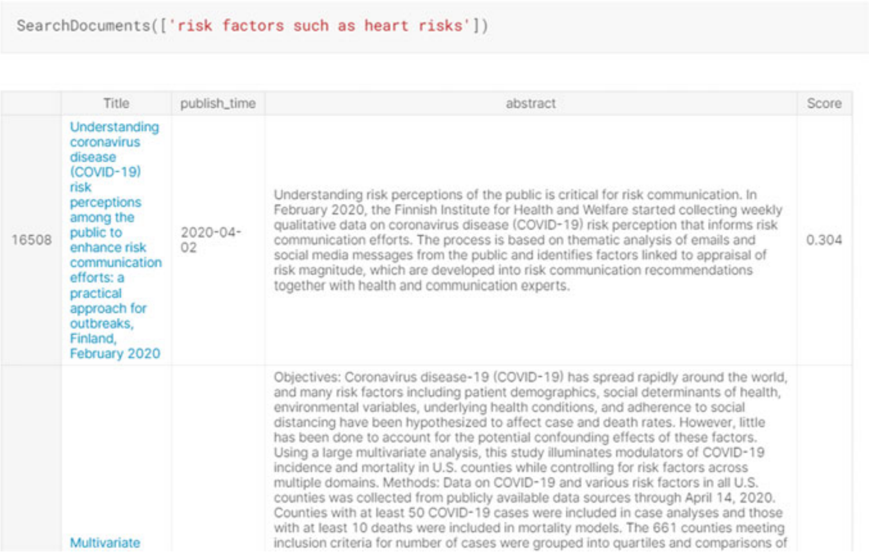
Fig. 1 Finding documents related to COVID-19 risk factors



```
SearchDocuments(['Risk factors such as Smoking, pre-existing pulmonary disease'])
```

	Title	publish_time	abstract	Score
40112	Smoking is Associated with COVID-19 Progression: A Meta-Analysis	2020-04-16	Objective: To determine the association between smoking and progression of COVID-19. Design: A meta-analysis of 12 published papers. Data Source: PubMed database was searched on April 6, 2020. Eligibility criteria and data analysis: We included studies reporting smoking behavior of COVID-19 patients and progression of disease. Search terms included smoking, smoker*, characteristics, risk factors, outcomes, and COVID-19, COVID, coronavirus, sar cov-2, sar cov 2. There were no language limitations. One author extracted information for each study, screened the abstract or the full text, with questions resolved through discussion among both authors. A random effects meta-analysis was applied. Main Outcome Measures: The study outcome was progression of COVID-19 among people who already had the disease. Results: We identified 12 papers with a total of 9,025 COVID-19 patients, 878 (9.7%) with severe disease and 495 with a history of smoking (5.5%). The meta-analysis showed a significant association between smoking and progression of COVID-19 (OR 2.25, 95% CI 1.49-3.39, p=0.001). Limitations in the 12 papers suggest that the actual risk of smoking may be higher. Conclusions: Smoking is a risk factor for progression of COVID-19, with smokers having higher odds of COVID-19 progression than never smokers. Physicians and public health professionals should collect data on smoking as part of clinical management and add smoking cessation to the list of practices to blunt the COVID-19 pandemic.	0.428
16516	Smoking Upregulates Angiotensin-Converting Enzyme-2 Receptor: A Potential	2020-03-20	The epicenter of the original outbreak in China has high male smoking rates of around 50%, and early reported death rates have an emphasis on older males, therefore the likelihood of smokers being overrepresented in fatalities is high. In Iran, China, Italy, and South Korea, female smoking rates are much lower than males. Fewer females have contracted the virus. If this analysis is correct, then Indonesia would be expected to begin experiencing high rates of Covid-19 because its male smoking rate is over 60% (Tobacco Atlas). Smokers are vulnerable to respiratory viruses. Smoking can upregulate angiotensin-converting enzyme-2 (ACE2) receptor, the known receptor for both the severe acute respiratory syndrome (SARS)-coronavirus (SARS-CoV) and the human respiratory	0.332

Fig. 2 Finding documents related to risk factors like smoking and pre-existing pulmonary disease



```
SearchDocuments(['risk factors such as heart risks'])
```

	Title	publish_time	abstract	Score
16508	Understanding coronavirus disease (COVID-19) risk perceptions among the public to enhance risk communication efforts: a practical approach for outbreaks, Finland, February 2020	2020-04-02	Understanding risk perceptions of the public is critical for risk communication. In February 2020, the Finnish Institute for Health and Welfare started collecting weekly qualitative data on coronavirus disease (COVID-19) risk perception that informs risk communication efforts. The process is based on thematic analysis of emails and social media messages from the public and identifies factors linked to appraisal of risk magnitude, which are developed into risk communication recommendations together with health and communication experts.	0.304
	Multivariate		Objectives: Coronavirus disease-19 (COVID-19) has spread rapidly around the world, and many risk factors including patient demographics, social determinants of health, environmental variables, underlying health conditions, and adherence to social distancing have been hypothesized to affect case and death rates. However, little has been done to account for the potential confounding effects of these factors. Using a large multivariate analysis, this study illuminates modulators of COVID-19 incidence and mortality in U.S. counties while controlling for risk factors across multiple domains. Methods: Data on COVID-19 and various risk factors in all U.S. counties was collected from publicly available data sources through April 14, 2020. Counties with at least 50 COVID-19 cases were included in case analyses and those with at least 10 deaths were included in mortality models. The 661 counties meeting inclusion criteria for number of cases were grouped into quartiles and comparisons of	

Fig. 3 Finding documents related to risk factors like heart risks

5 Conclusion

As new coronavirus empirical research expands swiftly, the medical analysis organization finds it extraordinarily difficult to follow up on current alerts. Extraction of efficient and relevant patient-centered COVID-19 risk factors information using artificial intelligence. This provides insights over the ongoing battle against this infectious disease from respective natural language processing tools. Consequently, the medical research communities are in great demand for these methods and make use of linguistic properties. This article follows a robust approach as a search process from a given input keyword. This creates and supports a well-defined framework as the platform includes extractive summaries (SciBert). In a fraction of the time, the researchers were able to read the article easily. Furthermore, this framework can turn into an expert classification framework with a little creativity, where specialists can asynchronously click on the correct fragment phrases to transform an unsupervised approach to a supervised learning function.

Acknowledgements This research article work does not have any funding support from any funding agencies.

References

1. World Health Organization (2020) Cumulative number of reported probable cases of severe acute respiratory syndrome (SARS). https://www.who.int/csr/sars/country/en
2. He JF et al (2004) Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. Science 303:1666–1669
3. World Health Organization—WHO (2019) Middle East respiratory syndrome coronavirus (MERS-CoV). https://www.who.int/emergencies/mers-cov/en
4. Alserehi H, Wali G, Alshukairi A, Alraddadi B (2016) Impact of middle east respiratory syndrome coronavirus (MERS-CoV) on pregnancy and perinatal outcome. BMC Infect Dis 16:105
5. Kumar P, Medigeshi GR, Mishra VS, Islam M, Randev S, Mukherjee A, Chaudhry R, Kapil A, Ram Jat K, Lodha R, Kabra SK (2017) Etiology of acute respiratory infections in infants: a prospective birth cohort study. Pediatr Infect Dis J. 36(1):25–30
6. Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T (2006) Bats: important reservoir hosts of emerging viruses. Clin Microbiol Rev 19(3):531–545
7. Boro AR, Saikia PK, Saikia U (2018) New records of bats (Mammalia: Chiroptera) from Assam, northeastern India with a distribution list of bat fauna of the state. J Threatened Taxa 10(5):11606–11612
8. Kumar NP, Shankarego HS, Revathy R (2011) An assessment of preventable risk factors for chronic non-communicable diseases in an adult population. Asian J Epidemiol 4(1):9–16
9. Lim SS, Vos T, Flaxman AD et al (2012) Comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. 2e Lancet 380(9859):2224–2260
10. Institute for Health Metrics and Evaluation (2014) India high blood pressure. https://www.Healthmetricsandevaluation.org/search-gbd-data
11. Ganesh K, Naresh A, Bammigatti C (2015) Prevalence and risk factors of hypertension among male police personnel in urban Pondicherry, India. Kathmandu Univ Med J 12(4):242–246
12. Bartwal J, Awasthi S, Rawat CMS, Singh RK (2014) Prevalence of hypertension and its risk factors among individuals attending outpatient department of rural health training centre, Haldwani. Ind J Commun Health 26(1):76–81
13. Kokiwar PR, Gupta SS, Durge PM (2012) Prevalence of hypertension in a rural community of central India. J Assoc Physicians India 60:26–29
14. Rajkumar E, Romate J, Factors BR (2020) Hypertension knowledge, and hypertension in rural India, Hindawi. Int J Hypertens. 1–7 Article ID 8108202. https://doi.org/https://doi.org/10.1155/2020/8108202
15. World Health Organization (2005) 2e WHO step wise approach to chronic disease risk factor surveillance. World Health Organization, Geneva, Switzerland
16. Dhama K, Khan S, Tiwari R, Dadar M, Malik Y, Singh K, Chaicumpa W (2020) COVID-19, an emerging coronavirus infection: advances and prospects in designing and developing vaccines, immunotherapeutics and therapeutics. Hum Vaccines Immunother doi: https://doi.org/10.1080/21645515.2020.1735227
17. Chobanian AV, Bakris GL, Black HR et al (2003) Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure. Hypertension 42(6):1206–1252
18. Ford N, Vitoria M, Rangaraj A, Norris SL, Calmy A, Doherty M (2020) Systematic review of the efficacy and safety of antiretroviral drugs against SARS, MERS or COVID-19: initial assessment. J Int AIDS Soc 23(4):e25489. https://doi.org/10.1002/jia2.25489
19. Tu YF, Chien CS, Yarmishyn AA et al (2020) A review of SARS-CoV-2 and the ongoing clinical trials. Int J Mol Sci 21(7):2657. doi:https://doi.org/10.3390/ijms21072657
20. Yang Y, Islam MS, Wang J, Li Y, Chen X (2020) Traditional Chinese medicine in the treatment of patients infected with 2019-new coronavirus (SARS-CoV-2): a review and perspective. Int J Biol Sci 16(10):1708–1717. Published 2020 Mar 15. doi:https://doi.org/10.7150/ijbs.45538

21. Russell B, Moss C, George G et al (2020) Associations between immune-suppressive and stimulating drugs and novel COVID-19-a systematic review of current evidence. Ecancermedicalscience 14:1022. Published 2020 Mar 27. doi:https://doi.org/10.3332/ecancer.2020.1022
22. Ruocco G, Feola M, Palazzuoli A (2020) Hypertension prevalence in human coronavirus disease: the role of ACE system in infection spread and severity. Int J Infect Dis 95:373–375
23. Kapoor KM, Kapoor A (2020) Role of chloroquine and hydroxychloroquine in the treatment of COVID-19 infection. Syst Lit Rev. medRxiv 2020.03.24.20042366; doi: https://doi.org/https://doi.org/10.1101/2020.03.24.20042366
24. Jeria RB, Reyes MX, Franco JV, Acuna MP, Torres Lopez LA, Rada G (2020) Chloroquine and hydroxychloroquine for the treatment of COVID-19: a living systematic review protocol. medRxiv 2020.04.03.20052530; doi: https://doi.org/https://doi.org/10.1101/2020.04.03.20052530
25. Canrong Wu, Liu Y, Yang Y, Peng Zhang Wu, Zhong YW, Wang Q, Yang Xu, Li M, Li X, Zheng M, Chen L, Li H (2020) Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. Acta Pharmaceutica Sinica B 10(5):766–788

Design and Development of Topic Identification Using Latent Dirichlet Allocation

P. Lakshmi Prasanna, S. Sandeep, V. Kantha Rao, and B. Sekhar Babu

Abstract Data storing and retrieving are the most important task in the present condition. Storing can be ended based on the topic that the document describes. Text mining generates documents from the collection of topics. To identify the topics, we have to categorize the documents; to classify, we are using topic modeling. Text mining technique is used for discovering latent semantic structure which is a fragment of topic modeling. Various research areas that make use of probabilistic modeling includes software engineering, political science, and medical science. A topic model is a probability-based model that discovers the major themes which are a group of documents. The main idea is to treat the documents as mixtures of topics in the topic model, and every topic is viewed as a probability distribution of the words. This research work aims to propose a model called topic modeling using LDA, and this model has been experimented on two datasets, where one is two news group dataset, and other is twenty news group dataset, and finally, all the results are tabulated.

1 Introduction

LDA is an unconventional probably based model for compilations of detached information and consequently more suitable for document data. An unsubstantiated accession was supposedly to be utilized for identifying and penetrating the bunch of terms in huge clusters of documents. This method imagines that each text is a combination of concepts; each and every word is accredited to each of the concepts that has limited prospect. This model also uncovers various topics that the texts correspond to and to what extent every individual concept is at hand in a document. LDA bears to discover the credibility dispersions over terms; then, it finds arrays of words that form together with definite possibility. Aforementioned arrays are marked as “topic.”

LDA is a Bayesian inference model designed by David et al. [1]. Every text is linked to a likelihood allocation more than topics, and further, topics are probability

---

P. Lakshmi Prasanna (✉) · S. Sandeep · V. Kantha Rao · B. Sekhar Babu
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India
e-mail: pprasanna@kluniversity.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_63

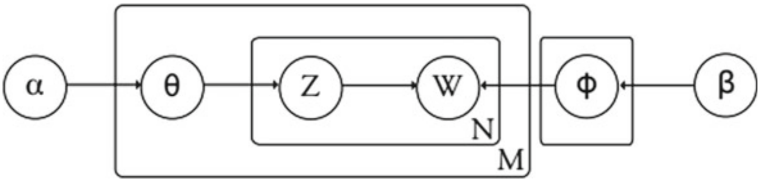


Fig. 1 Block diagram of LDA

segregation son terms. Every expected result of an arbitrary is changeable with the possibility that the happening will occur and come out of the distribution as an equation that links them. For example, in a sport coin flips for decision of who will go first and second as a deciding factor, there are two probable results out of it: heads or tails. Heads is personified as 1 and tails as 0. Here is an indiscriminate variable which is marked X. And dualistic expected reactions is delineated by X as 0 or X as 1. The probability allotment of X or $P(X = x)$ is: $P(X = 0) = 0.5$ $P(X = 1) = 0.5$. Figure 1 represents the block diagram of LDA.

2 LDA Criterion

To attain rest results, the below mentioned few parameters can be helpful.

1. Topics Number—Quantity of topics which have to be acquired through the body.
2. α and β are hyper parameters—α symbolizes document-topic density, and β symbolizes topic and token density. Depending upon the upper worth of α, texts are possessed of added topics and greater the β; themes are made up of bigger pack of terms [2].
3. Iterations no—The duration for algorithm to come together (or how do we sense that the algorithm has converged)?

1. The hidden topics digit in the corpus (K) needs to be explained previous to initiating the practice of the model.
2. Since the perplexity reduces very slowly, algorithm convergence needs a large number of iterations.
3. If the training information set is large, then the algorithm requires a long time for preparation.
4. Being a probabilistic model, LDA requires additional clarifications in preparing statistical inference. Hence, it does not toil on same small texts like sentence categorization or tweets.
5. For providing meaningful knowledge to advance excellence of topics, LDA does not take into account connection of concepts or words.

3 Topic

To identify what is a topic and how it can be interpreted has been a continuous point of discuss in the literature. Commonly, themes provide an outline to the subject matter present in the texts being considered as well as brief explanation of the contents of the dataset. Whether the statement "topics" is even the appropriate word to use for the groups of words generated by the topic model evaluation which has been under radar for decades. "Topic" might not constantly be apt word to be use for the divisions which are generated through topic modeling [3]. Some have recommended that "discourse" may be sophisticated, as subjects are not constantly amalgamated linguistically. For instance, mathematics, lineage, reasoning, science, theory, century, and epistemology shall not be well thought-out semantically coherent as they are seen as discourse of subjects found contained by philosophy. When every word has a different meaning, it would be intricate to give them an explicit topic [4].

3.1 Topic Models

Topic models are a class of algorithms which exploit co-occurrences of words in documents in order to discover hidden sets of words which explain the co-occurrence patterns and are referred to as topics. Probabilistic topic models explain pragmatic documents with an underlying, hidden probabilistic model. The observed documents are assumed to be random samples from this model. In a probabilistic topic model, each document is associated in the midst of a probability allocation over a set of ideas, and concepts get associated with a prospective disposal upon the set of words. Similar documents share a similar topic distribution. Topic models are often employed in regard to text mining errands, e.g., to have understanding and visualizing the content of large document corpora or for detecting relations between topics and other variables of interest [5]. Additionally, topic models can be employed as a mean for dimensionality reduction (documents are mapped to a lower dimensional topic space), as input for prediction tasks, in recommender systems (e.g., for predicting semantically related tags) or in information retrieval (e.g., to understand and disambiguate the topic of query terms).

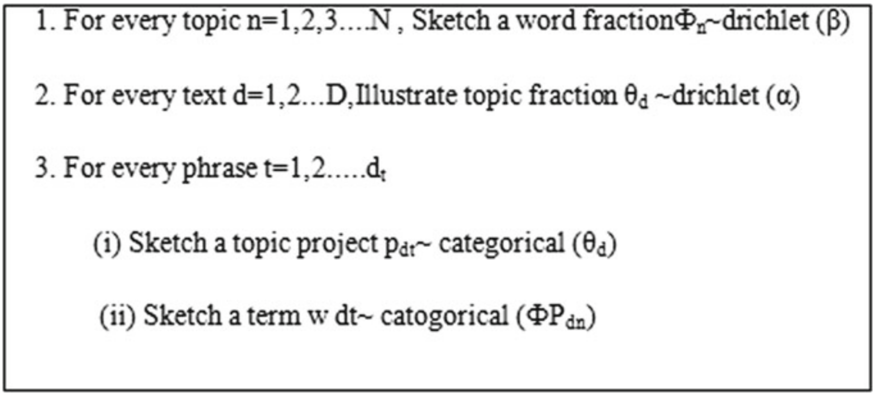
3.2 Topic Modeling Using LDA

For finding topics, LDA algorithm was used by applying Bayesian belief networks theorem to calculate probability and topic identification of each term and each document. For feature reductions, LDA Algorithm was considered to reduce features and filtered top terms of each topic and store it in the filtered document term matrix.

Latent Dirichlet allocation is being primarily utilized for evaluating texts [2]. It presumes that here is N no. of topics on how texts will be produced, and each topic is corresponding to multinomial circulation on words in the terminology. A document $w_{\mathrm{d}} = \{w_{\mathrm{dt}}\}d_{\mathrm{tt}} = 1$ is accomplished by variety a concoction and these topics and fragmenting words from the mixture [1]. Figure 2 indicates the procedure of LDA Algorithm, and Table 1 represents the notations of LDA.

Applied the Bayesian probability to identify the model.

- Represent β as the total terms of the documents.
- Represent θ as specifying all the topics.
- Display all the terms in the documents with a minimum range of 10.
- Dividing the terms into topics.
- All terms are displayed in the topic wise.
- Display all the probability values with words.
- Based on these probabilities to create a word cloud.



1. For every topic n=1,2,3….N , Sketch a word fractionΦ_n~drichlet (β)

2. For every text d=1,2…D,Illustrate topic fraction θ_d ~drichlet (α)

3. For every phrase t=1,2…..d_t

    (i) Sketch a topic project p_{dt}~ categorical (θ_d)

    (ii) Sketch a term w dt~ catogorical (ΦP_{dn})

Fig. 2 Procedure of LDA Algorithm

Table 1 Notations of LDA

Symbol	Description
N	No. of topics
Y	No. of unique words in the vocabulary
D	No. of documents
DT	No. of words in the documents DT
θ_{d}	Proportion of topics specific to documents
On	Proportion of words specific to topic N
P_{dn}	Identify the topics of *n*th word in document d
W_{dt}	Identify the word in document d
αβ	Parameters of Dirichlet distribution

4 Procedure of LDA

See Fig. 2.

5 Notations of LDA Algorithm

Here, the first step depicts the number of topics and after that represents all word momentarily apportion toward the topics, and the procedure is completed absurdly, and from time to time, similar terms can be functional to various topics. The final step shows the updated adaptation of the topic assignment depended on their credibility as per the above criteria:

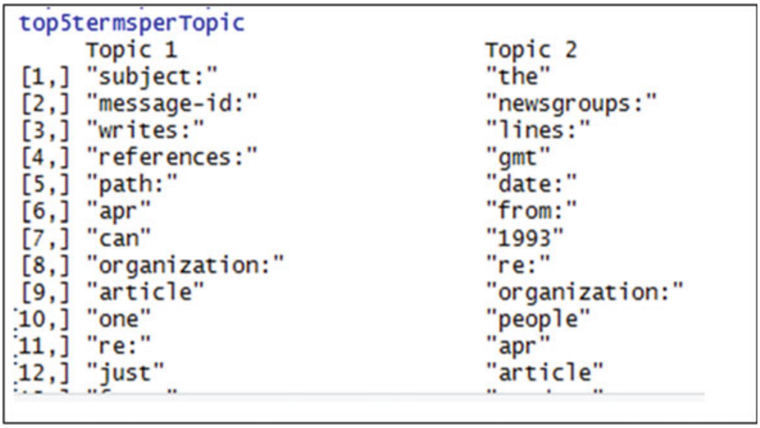
1. Primary criteria is about the length of prevalence is that tokens across the topics—it can be called as $P(w/t)$.
2. The another criteria is about the topic of the issues in the document $P(t/d)$.

As per the Bayesian belief network theorem $P(t/w) = P(t/d) * P(w/t)$. To calculate probability of each term. Figure 3, 4, 5, 6, and 7 represent the outcome of topic modeling using LDA.

Top terms per topics for 2 groups data (Figs. 3 and 4).

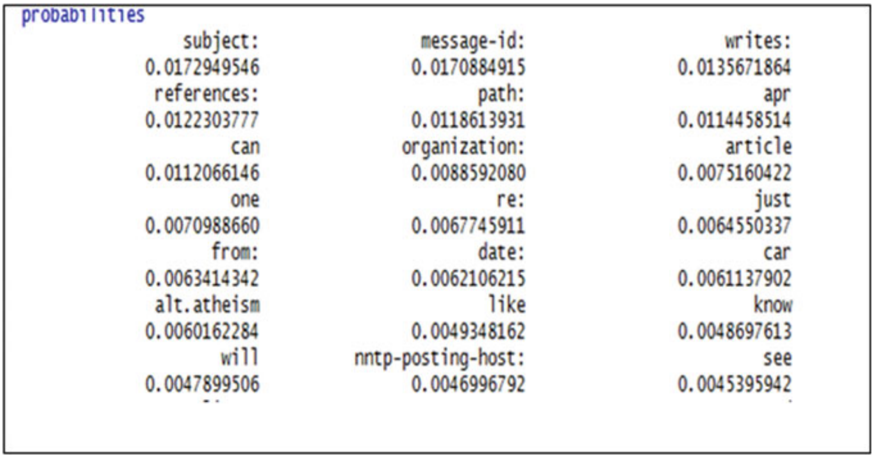
Top terms with probabilities for 20 news groups data (Figs. 5 and 6).

Filtered document term matrix (FDTM) (Fig. 7).



```
top5termsperTopic
      Topic 1                        Topic 2
 [1,] "subject:"                     "the"
 [2,] "message-id:"                  "newsgroups:"
 [3,] "writes:"                      "lines:"
 [4,] "references:"                  "gmt"
 [5,] "path:"                        "date:"
 [6,] "apr"                          "from:"
 [7,] "can"                          "1993"
 [8,] "organization:"                "re:"
 [9,] "article"                      "organization:"
 10,] "one"                          "people"
 11,] "re:"                          "apr"
 12,] "just"                         "article"
```

Fig. 3 Top terms per topics for 2 groups data

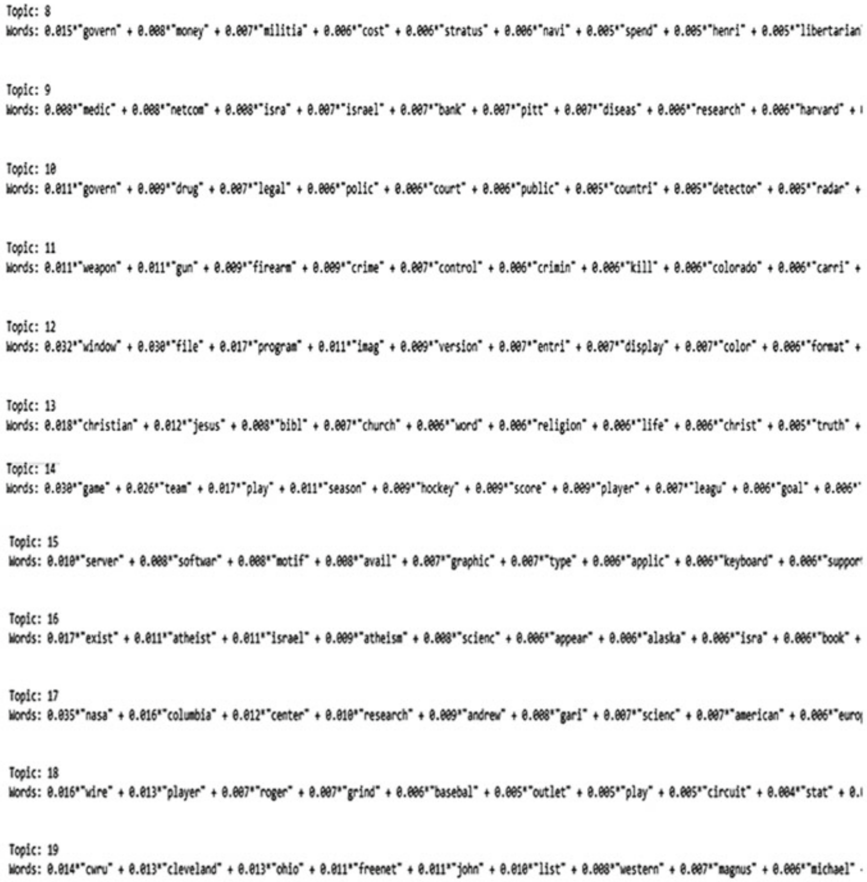


```
probabilities
          subject:             message-id:               writes:
      0.0172949546            0.0170884915          0.0135671864
       references:                   path:                   apr
      0.0122303777            0.0118613931          0.0114458514
               can           organization:               article
      0.0112066146            0.0088592080          0.0075160422
               one                     re:                  just
      0.0070988660            0.0067745911          0.0064550337
             from:                   date:                   car
      0.0063414342            0.0062106215          0.0061137902
       alt.atheism                    like                  know
      0.0060162284            0.0049348162          0.0048697613
              will      nntp-posting-host:                   see
      0.0047899506            0.0046996792          0.0045395942
```

Fig. 4 Top terms with probabilities for 2 groups data

6 Conclusion

Topic modeling began as of text mining method designed for disclosing latent semantic structure within a compilation of texts. In text mining, each archive is produced from anthology of topics. It relies upon probabilistic modeling that has a huge assortment of relevance such as image detection, semantic understanding, and automatic music improvisation recognition. In this chapter to proposed topic modeling employing latent Dirichlet allocation [LDA], the LDA works backward to learn the topic illustration in all texts and the word allotment to every topic. The main focus of this paper is on LDA algorithms, and the outcomes will be displayed in 20 news group dataset and 2 groups dataset.



```
Topic: 8
Words: 0.015*"govern" + 0.008*"money" + 0.007*"militia" + 0.006*"cost" + 0.006*"stratus" + 0.006*"navi" + 0.005*"spend" + 0.005*"henri" + 0.005*"libertarian"

Topic: 9
Words: 0.008*"medic" + 0.008*"netcom" + 0.008*"isra" + 0.007*"israel" + 0.007*"bank" + 0.007*"pitt" + 0.007*"diseas" + 0.006*"research" + 0.006*"harvard" +

Topic: 10
Words: 0.011*"govern" + 0.009*"drug" + 0.007*"legal" + 0.006*"polic" + 0.006*"court" + 0.006*"public" + 0.005*"countri" + 0.005*"detector" + 0.005*"radar" +

Topic: 11
Words: 0.011*"weapon" + 0.011*"gun" + 0.009*"firearm" + 0.009*"crime" + 0.007*"control" + 0.006*"crimin" + 0.006*"kill" + 0.006*"colorado" + 0.006*"carri" +

Topic: 12
Words: 0.032*"window" + 0.030*"file" + 0.017*"program" + 0.011*"imag" + 0.009*"version" + 0.007*"entri" + 0.007*"display" + 0.007*"color" + 0.006*"format" +

Topic: 13
Words: 0.018*"christian" + 0.012*"jesus" + 0.008*"bibl" + 0.007*"church" + 0.006*"word" + 0.006*"religion" + 0.006*"life" + 0.006*"christ" + 0.005*"truth" +
Topic: 14
Words: 0.030*"game" + 0.026*"team" + 0.017*"play" + 0.011*"season" + 0.009*"hockey" + 0.009*"score" + 0.009*"player" + 0.007*"leagu" + 0.006*"goal" + 0.006*"

Topic: 15
Words: 0.010*"server" + 0.008*"softwar" + 0.008*"motif" + 0.008*"avail" + 0.007*"graphic" + 0.007*"type" + 0.006*"applic" + 0.006*"keyboard" + 0.006*"suppor

Topic: 16
Words: 0.017*"exist" + 0.011*"atheist" + 0.011*"israel" + 0.009*"atheism" + 0.008*"scienc" + 0.006*"appear" + 0.006*"alaska" + 0.006*"isra" + 0.006*"book" +

Topic: 17
Words: 0.035*"nasa" + 0.016*"columbia" + 0.012*"center" + 0.010*"research" + 0.009*"andrew" + 0.008*"gari" + 0.007*"scienc" + 0.007*"american" + 0.006*"euro

Topic: 18
Words: 0.016*"wire" + 0.013*"player" + 0.007*"roger" + 0.007*"grind" + 0.006*"basebal" + 0.005*"outlet" + 0.005*"play" + 0.005*"circuit" + 0.004*"stat" + 0.

Topic: 19
Words: 0.014*"cwru" + 0.013*"cleveland" + 0.013*"ohio" + 0.011*"freenet" + 0.011*"john" + 0.010*"list" + 0.008*"western" + 0.007*"magnus" + 0.006*"michael"
```

Fig. 5 Top terms with probabilities for 20 news groups data

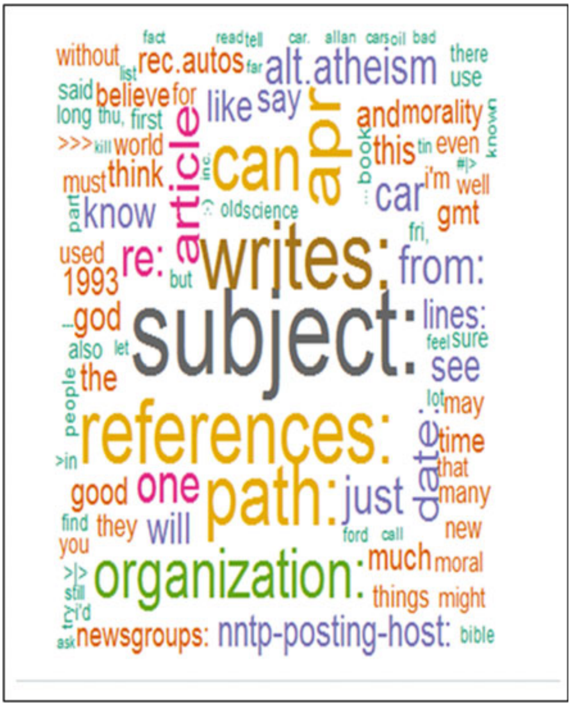
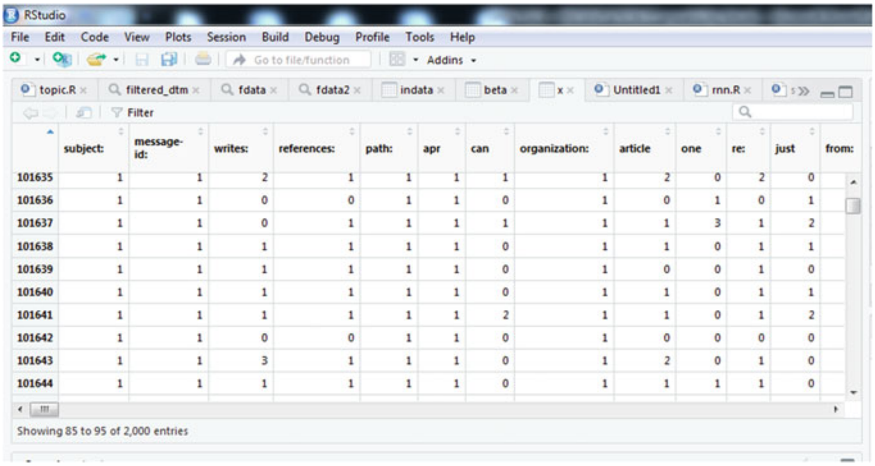


Fig. 6 News groups dataset of word cloud



	subject:	message-id:	writes:	references:	path:	apr	can	organization:	article	one	re:	just	from:
101635	1	1	2	1	1	1	1	1	2	0	2	0	
101636	1	1	0	0	1	1	0	1	0	1	0	1	
101637	1	1	0	1	1	1	1	1	1	3	1	2	
101638	1	1	1	1	1	1	0	1	1	0	1	1	
101639	1	1	1	1	1	1	0	1	0	0	1	0	
101640	1	1	1	1	1	1	0	1	1	0	1	1	
101641	1	1	1	1	1	1	2	1	1	0	1	2	
101642	1	1	0	0	1	1	0	1	0	0	0	0	
101643	1	1	3	1	1	1	0	1	2	0	1	0	
101644	1	1	1	1	1	1	0	1	1	1	1	0	

Showing 85 to 95 of 2,000 entries

Fig. 7 Filtered document term m0atrix for filtered features

References

1. BleiDM, Andrew Y (2003) Latent Drichlent allocation. J Mach Learn Res
2. Tong Z, Zhang H (2016) A text mining research based on Lda topic modelling. In: The sixth international conference on computer science, engineering and information technology
3. DayaSagar KV, Shyam Krishna C, Lalith Kumar G, Surya Teja P, Charless Babu G (2018) A method for finding threated web sites through crime data mining and sentiment analysis. Int J EngTechnol (UAE) 7(2):62–65
4. Wallach HM (2008) Structured topic models for language, PhD thesis
5. Roose H, Roose W, Daenekindt S (2018) Trends in contemporary art discourse: using topic models to analyze 25 years of professional art criticism. CulturSociol 12:303–324
6. Kousar A, Subrahmanyam K (2019) Feature selection, optimization and clustering strategies of text documents. Int J Electr Comput Eng 9(2):1313–1320
7. BleiDM (2012) Surveying a suite of algorithms that offer a solution to managing large document archives. Commun ACM
8. Bastani1 K, Namavari1 H, Shaffer J (2016) Latent Dirichlet Allocation (LDA) for topic modeling of the CFPB consumer complaints. IEEE
9. Kaur PC, Ghorpade T, RamraoAdik V (2017) Extraction of unigram and bigram topic list by using Latent Dirichlet Markov allocation and sentiment classification. In: 2017 international conference on energy, communication, data analytics and soft computing
10. PotharajuSP, Sreedevi M, AndeVK, TirandasuRK (2019) Data mining approach for accelerating the classification accuracy of cardiotocography. ClinEpidemiol Global Health
11. Poornima BK, Deenadayalan D, Kangaiammal A (2017) Text preprocessing on extracted text from audio/video using R. Int J Comput Intell Inform 6(4)
12. Sajid A, Jan S et al (2017) Automatic topic modeling for single document short texts. In: International conference on frontiers of information technology (FIT).
13. Sleeman J, Halem M, Finin T (2017) Discovering scientific influence using cross-domain dynamic topic modeling. In: 2017 IEEE international conference on big data (big data)
14. Sharma N, Yalla P (2017) Classifying natural language text as controlled and uncontrolled for UML diagrams. Int J Adv Comput Sci Appl
15. Sapul MSC, Aung TH, Jiamthapthaksin R (2017) Trending topic discovery of Twitter Tweets using clustering and topic modeling algorithms. In: 2017 14th international joint conference on computer science and software engineering (JCSSE)
16. Lakshmi Prasanna P, Rajeswara Rao D (2017) Literature survey on text classification: a review.J Adv Res Dyn Control Syst 9(12):2270–2280

Hand Gesture Controlled Robot

N. H. Prasad, A. Mariyan Richard, and B. N. Lakshmi Narayan

Abstract Due to the recent advancements in digital technologies, every human being has started to believe more in technology, where internet of things [IoT] assists people to remain more accurate in their work. So, IoT controlled car is a robot, which needs to be controlled by using human gestures. The user must wear a gesture device during which the sensor can be included. The hand movement will be recorded by the sensor for a specific direction, and it can end in deploying robotic motion within the respective directions. The robot and gesture instrument are connected wirelessly with help of radio waves. Wireless communication can help to interact with the robot in a more user-friendly way. In this, it commands the car by using accelerometer sensors that are connected to a hand glove. The sensors are intended to exchange the remote, which is generally utilized to run the car. This will allow the user to regulate actions, i.e., forward, backward, leftward, and rightward movements, while using an equivalent accelerometer sensor to regulate the throttle of the car.

Keywords Hand gestures · Accelerometer · Microcontroller · RF transceiver

1 Introduction

Nowadays, robotics is emerging as a ubiquitous paradigm field of technology. Robot is operated with the help of different system generate programs. An automatic robot does not require any human-driven control. Robot takes its own decision by sensing its environment. Most of the machine driven robots are automatic as they possess excess speed and great accuracy in their work. But some projects demand semi-automatic system or human controlled robots. Most of the control systems are based on gesture recognition or with voice and motion control. Only a small transmitting device should be worn in the palm, where it includes an accelerometer. This will

---

N. H. Prasad · A. Mariyan Richard (✉) · B. N. Lakshmi Narayan
Department of MCA, NMIT, Bengaluru, India
e-mail: mariyanrich01@gmail.com

B. N. Lakshmi Narayan
e-mail: narayan614@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_64

transmit an appropriate command to the robot and that it can perform any desired action. The data is then incorporated by a MCU and eventually our motor driver IC to regulate the motors.

The proposed research work is all about a robotic car, which is measured using hand gestures, i.e., the handling and controlling of the car depends on the gesture of user. In this project, gestures are captured by using accelerometer and it is computed by software namely, microcontroller software, and therefore, the specifications are forwarded to MCU and encoder circuit, where it is further processed by Nrf24L01 transceiver [1]. In the transceiver section, the Nrf24L01 transceiver is used to send and receive data by using radio waves and process it with MCU for providing those specifications to the robotic vehicles, which acts accordingly to the gesture.

1.1 Proposed Work

In our project, the hand gesture system includes two parts, namely transmitter and receiver. The transmitter part includes accelerometer, one RF-transmitter-module, and Arduino-nano board. The receiver part includes one RF-receiver-module, wheels, and motor driver IC's. The proposed project require two separate 5 V power supply, which can be applied to both the sections. This research work has three axis accelerometer but only two axis; i.e., x-axis and y-axis are used. The particular analog value is then converted to digital value. Then, the digital values are collected and processed by the Arduino nano and forwarded to the RF transmitter. Further, it will be received by the receiver and sends to the acceptor to stimulate the motor in a specific direction. When there is a tilt in the palm, the robotic car moves accordingly [2]. Hand gesture-based scars are controlled by using hand instead of buttons or joystick. Here, only hand movement is important to manage the robot (Fig. 1).

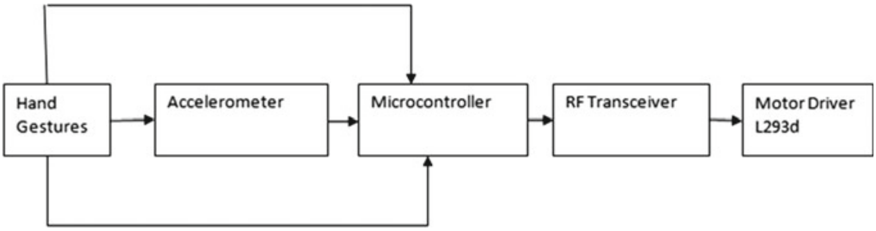


Fig. 1 Process of human-hand gestures

1.2 Block Diagram

1.2.1 Proposed Algorithm

1. Sensed the movement from hand to microcontroller by hand gesture.
2. MCU receives the data and will give the robot instructions.
3. Submit instructions to IC encoder.
4. Encoded data is transmitted over the transmitter.
5. The receiver receives the encoded data in the receivers end.
6. Receiver must forward the encoded data to the decoder.
7. Decoder decodes the data and then sends it to the driver.
8. In all actions, the motor driver drives the engine by following orders and gestures.
9. The robot eventually goes along with the movements.

2 Literature Survey

We described and discussed current work regarding gesture control robot in this segment. Hand-held interface for robot navigation. Using his or her hand movements, a robot can be operated by the user. To record the hand trajectories of a user, a 3-axis accelerometer is adopted. The data from the trajectory is transmitted wirelessly to a computer via an RF module. The trajectories obtained are divided into six control commands used to operate a robot. The classifier adopts the algorithm of dynamic time-warping to classify the hand trajectories. The current study also has drawbacks that the findings of simulation indicate that the classifier could only reach the right rate of 92.2%.

P. Davis et al. Throughout his work on visual gesture recognition, al shows the technique of using a model- based approach to understand human-hand gestures. This employs a small state machine to shape four qualitatively distinct aspects of a general gesture. In certain pictures, fingertips are traced to measure the translating trajectories. Instead the trajectories are practiced to get the gesture's start and stop position. Management is represented as a vector chart, and then matched to save gesture vector models using vector displacement-based table lookup. Results showing seven movements found using pictures examined at 4 Hz on a SPARC-1No different hardware. The seven movements are acting agents for left, right, up, down, catch, rotate, and avoid acts [3].

Rogalla et al. in his paper the use of gesture and speech control to order a robot assistant still requires classical user interfaces to provide advice to a mobile robot assistant. Oral or motion orders achieve a more natural way of directing. In this article, we present new approaches and improvements to existing methods in use in our lab. Our goal is to communicate with a robot to facilitate robust performance of simple tasks using natural and direct communication techniques. In this paper, we explain the framework for understanding the robot's vision and voice. Then, we display robot control to select the appropriate robot response to solve basic manipulations [4].

In his paper, hand gesture recognition with depth images**, Jesus Suarez** et.al presents a review of the literature on the use of depth for hand tracking and recognition of gestures. P.M. In his paper, GESTURE CONTROL ROBOT, Singh et.al presumes that human–machine intercommunication is shifting away from mouse and pen today and is turning the mechanical environment into a widespread and much more cooperative one. With each day, the passage between computers and humans is reduced with the advent of new technology to reduce living standards. Maneuvers have played a critical role in depreciating this gulf. A dogmatic study of "Human–Machine Interaction" using gestures9 has been acted upon in this article. The architecture of the machine is divided into two parts namely: part of the accelerometer and part of the robot [4].

Rahul Ranjan Singh presents a wirelessly controlled robot on his paper, wireless Controlled robot movement device designed to prepare microcontroller, which can be operated human gestures by simple. Within the transmitter device, the robot may move forward, backwards, left, and right. The sensor will admit the hand movement in a particular direction resulting in the robot's motion in the particular direction. Half of the accelerometer and half of the robot. The wireless receiver module takes on the wave signal received by the joystick wireless transmitter module. The module then decodes the wave signal and sends it over UART to the ATmega328P microcontroller. UART looks for Universal asynchronous transmitter receivers. It is commonly used as a protocol for microcontroller interaction. The microcontroller then transmits the motor controlling signal according to the agreed wireless signal to the motor driver IC L298.

2.1 Proposed Methodology

Above figure shows that the gesture movements/direction is sensed from hand to microcontroller. So, the particular encoded values are used to encode the values. And are collected and processed through the nano board to the RF transmitter. Then, in the receiver end, the received data is transmitted to the decoder and stimulate motors to a particular direction. The robot moves in four directions namely backward, forward, right, and left. When there is tilt/bend in the palm, the robotic car moves accordingly. Hand gesture controlled cars are controlled by using hand instead of buttons or joystick. Here, only hand movement is important to manage the robot [5] (Fig. 2).

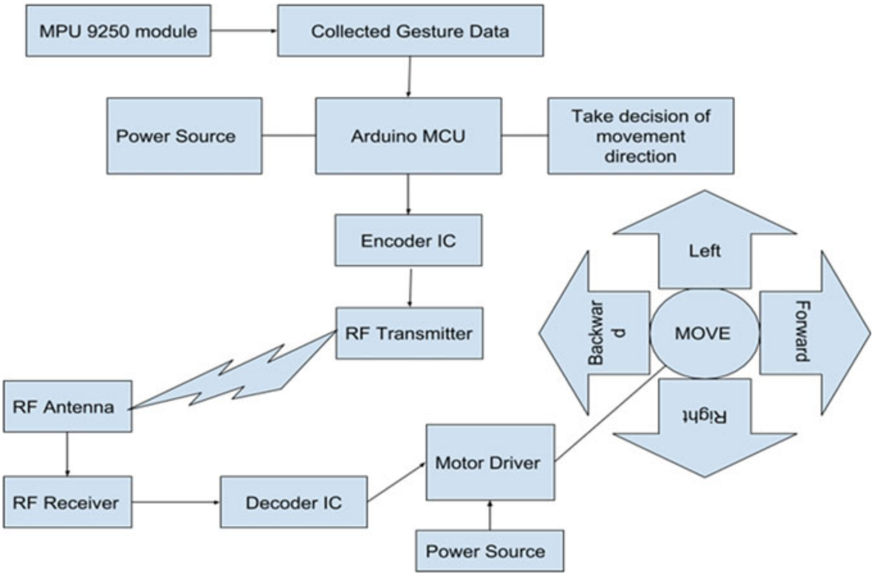
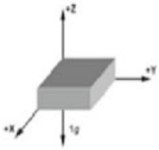
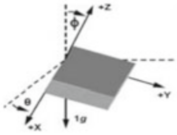
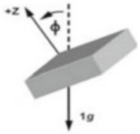
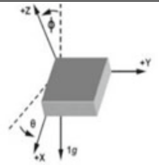
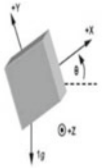
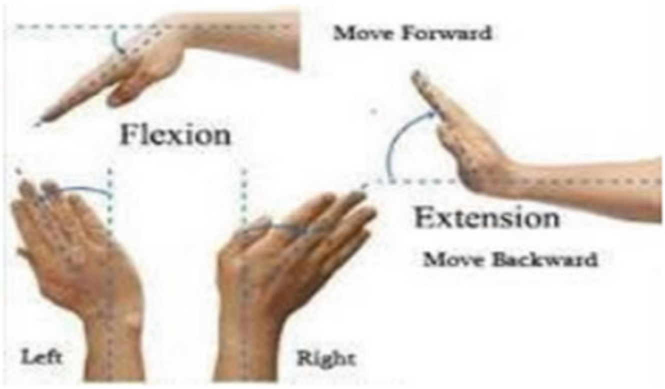


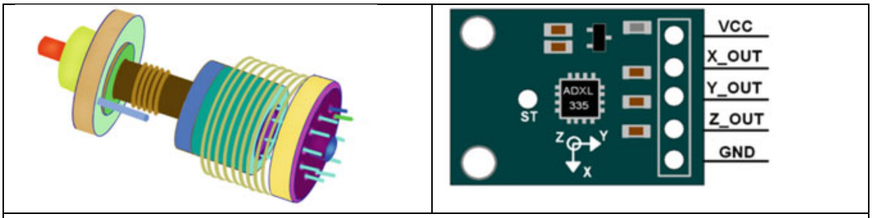
Fig. 2 Methods and process of hand gesture using Arduino

				
Stop condition	Forward Motion	Backward motion	Right Motion	Left Motion



Hand Gestures are

Accelerometer is an instrument that is used to measure the vibration or acceleration of the movement of a hand, where the force caused by vibration or change in motion that causes the mass to squeeze the piezoelectric material, which produces an electrical charge that is proportional to the force exerted upon it.

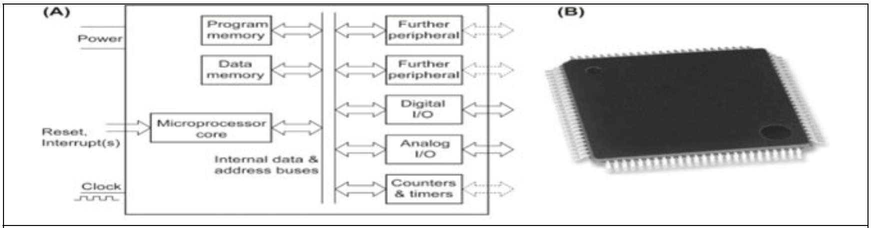


Accelerometer

2.2 *Components*

2.2.1 MCU Arduino Nano Board

Arduino-nano board is also called as MCU, and it is created by Arduino.cc. This Arduino-nano is used in Arduino-UNO board and designed with microcontroller like Atmega328. It is used all over world with variety of applications and specifications. An Arduino board also includes Arduino Mega, Arduino Pro Mini, Arduino UNO, Arduino YUN, Arduino Lily pad and Arduino Leonardo, etc. [6].



Microcontroller diagram

2.2.2 Module NRF24L01+RF

It is used all over the world for its wide variety of applications that require wireless control. It is also called as single chip radio transceiver for the worldwide with its respective 2.4–2.5 GHz ISM band. In this, every module transmits and receives data from each other. These modules are cheap/low cost and you can use them with any MCU. This transceiver consists of a completely integrated frequency synthesizer, modulator, demodulator crystal oscillator, and a protocol engine [7]. The RF module is used to transmit and receive the data between two devices through radio signals. It is also used in embedded system for wireless communication.

2.2.3 L298N Motor Driver

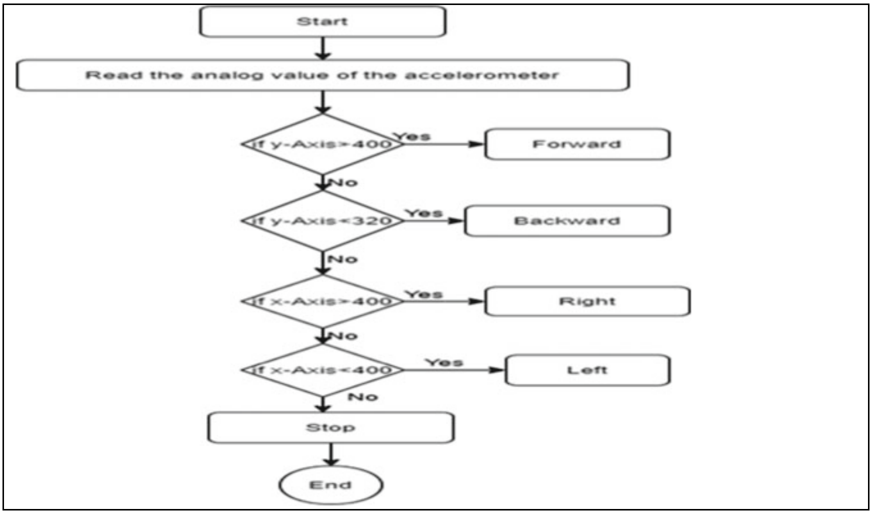
The L298N motor driver is used to permit speed and direction control to two DC motors at the same time. It is also called as dual H bridge motor driver. Which poses 2 terminal blocks for two motors namely motor A and motor B. This in turn has high voltages between DC motors.

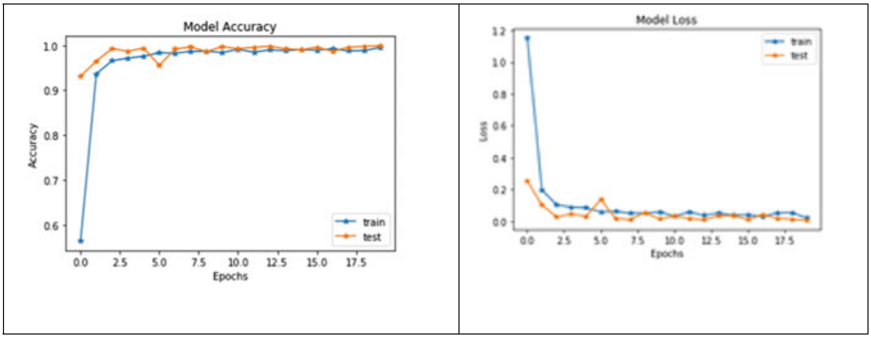
2.2.4 AdXl335

ADXL335 is a three axis accelerometer. This reads the *X*, *Y*, and *Z* axis of accelerometer values as analog voltages. It can measure the tilted values of accelerometer due to gravity. The accelerometer is used to sense the amount of speed and in which direction the device is moving.

2.2.5 TT Gear Motor

TT gear motor is also called as TT DC gearbox motor with great ratios in it. It is low cost plastic gearbox motor, and it is used to make our projects move. This TT DC gearbox motor provides gear ratio of 1:48, which can provide nearly double the speed of the blue metal TT DC motor that has 1:90 gear ratios. Overall it is used to control the speed via PWM from Arduino or any MCU [8].





3 Conclusion

This paper was successfully implemented and tested. The testing was done for test cases, and it was found that it worked well for all conditions. That follows my hand gesture and works according to the data which is transmitted via hand wirelessly. This device measures the gestures and transmits the particular instructions to the robot to maneuver consistently with the user.

4 Future Enhancements

The proposed system is applicable in dangerous environment where cameras are attached to the robot often and may be viewed by the user who is in his station. In medical field, where miniature robots are developed which will help doctors for efficient surgery operations, and it will be used. For more methodical response, threshold values are often employed to detect gesture and advanced features like finger counts that provide distinct functional commands can be used. We can add video camera on the internet for live streaming. We can add bomb and metal detectors and can dispatch to a place, which is harmful for an individual to travel. This sort of hand gesture system is often developed for whole body and can be utilized in military operations.

References

1. Taipei T (2010) The 2010 IEEE/RSJ international conference on intelligent robots and systems. October 1822
2. Kun, Miller WT (1996) Adaptive dynamic balance of abipedrobotusing neural networks. In: IEEE conference on robotics and automation, pp 240–245
3. Zhang S, Wang C, Chan S-C, Wei X, Check HH (2015) IEEE Sens J 15(5)
4. Flikkema P (2014) Department of Electrical Engineering and Computer Science, Northern Arizona University Flagstaff, Arizona
5. Wu Y, Lim J, Yang M-H (2013) A benchmark. In: Proceedings of IEEE conference computer vision and pattern recognition: online object tracking, pp 2411–2418
6. Zhabelova G, Vyatkin V: Multiagent smart grid automation architecture based on IEC 61850/61499 intelligent logical nodes. IEEE Trans Indus Electr 59(5):2351–2362
7. Suarez J, Murphy RR (2012) Hand gesture recognition with depth images: a review. 9–13 Sept 2012. doi: https://doi.org/10.1109/ROMAN.2012.6343787
8. Mansuri R, Vakale S, Shinde A, Patel T (2013) Hand gesture control robot vehicle. IJECT 4(SPL-2):77–80
9. Aswath S, Krishna Tilak C, Suresh A, Udupa G (2014) Human gesture recognition for real—time control of humanoid robot. Int J AdvMechAutomobEng (IJAMAE) 1(1):96–100
10. Jain M, Aditi AL, Fahad Khan M, Maurya A (2012) Wireless gesture control robot: an analysis. Int J Adv Res ComputCommunEng 1(10)
11. ShivrajSBN, Sumathi S (2014) Hand gesture based direction control of Robocar using Arduino microcontroller. Int J Recent TechnolEng (IJRTE) 3(3):32–35
12. Somkuwar V, SandeepkakdeRG, Design and implementation of gesture controlled robot using flex sensor and accelerometer
13. Baron G, Czekalski P, Golenia M, Tokarz K (2015) Gesture and voice driven mobile tribot robot using Kinect sensor. IEEE 33–34
14. Thi-Lan L, Minh-Quoc N, Thi-Thanh-Mai N (2013) Human posture recognition using human skeleton provided by Kinect. IEEE 340–345
15. Matthew T (2010) Recognition hand gesture with Microsoft's Kinect. In: Paper written for CS228 winter 2010
16. Jamie S, Andrew F, Mat C, Toby S, Mark F, Richard M, Alex K, Andrew B. Real-time human pose recognition in parts from single depth images. Microsoft Research Cambridge & Xbox Incubation

17. Lee J, Hironori T, Hitoshi Y, Akihiro K, Yasue M (2013) A robust gesture recognition based on depth data. In: The 19th Korean-Japan joint workshop on frontiers of computer vision, pp 127–131
18. Von Marina S (2013) A introduction to depth sensor for gesture recognition. Written for State University of New York at Stony Brook, pp 1–11
19. Mantecón T, del Blanco CR, Jaureguizar F, García N (2016) Hand gesture recognition using infrared imagery provided by leap motion controller. In: International of conference on advanced concepts for intelligent vision systems, ACIVS 2016, Lecce, Italy, pp 47–57, 24–27 Oct 2016. doi: https://doi.org/10.1007/978-3-319-48680-2_5
20. https://www.researchgate.net/publication/_Gesture_Controlled_Robot_using_Arduino_and_Android
21. https://www.ijitee.org/wp-content/uploads/papers/v9i2/B7185129219.pdf
22. https://www.elprocus.com/h-bridge-motor-control-circuit-using-l293d-ic/
23. www.ijsr.netvol 4, Issue 3, March 2015

Interdependence in Artificial Intelligence to Empower Worldwide COVID-19 Sensitivity

E. Laxmi Lydia, Jose Moses Gummadi, Chinmaya Ranjan Pattanaik, A. Krishna Mohan, G. Jaya Suma, and Ravuri Daniel

Abstract Researchers from different disciplines are striving to leverage a solution for COVID-19 with a unique commitment of scientific collaborations and with cognitive technologies, and highly flexible learning processes are required to maintain the transmission of knowledge, prototype, and code by integrating the application areas to a specific culture and cross-border cooperation. The research experts in the artificial intelligence (AI) and machine learning (ML) domain were tracked and predicted with real-time data observed throughout the world regarding the pandemic situation and timely assessment of the distributed COVID-19 patient information. The considered physiological features followed by clinical tests of patients with COVID-19 offer very simple access to subsequent data transformation, which was relevant but complicated. This paper works on in-depth exploratory data analysis (EDA) prediction analysis over the global medical database of COVID-19 will be available for benefiting future artificial predictive, analytical, and biomedical research, which includes additional COVID-19 approaches associated with pandemics.

E. Laxmi Lydia (✉)
Department of Computer Science and Engineering, Vignan's Institute of Information Technology (A), Visakhapatnam, Andhra Pradesh, India
e-mail: elaxmi2002@yahoo.com

J. Moses Gummadi
Department of CSE, VFSTR (Deemed to be University), Guntur, India
e-mail: josemoses@gmail.com

C. Ranjan Pattanaik
Department of Computer Science and Engineering, Ajay Binay Institute of Technology, Cuttack, Odisha, India
e-mail: chinmaya.pattnaik@rediffmail.com

A. Krishna Mohan
Department of Computer Science and Engineering, JNTUK, Kakinada, Andhra Pradesh, India

G. Jaya Suma
JNTUK-UCEV, Vizianagaram, India

R. Daniel
Department of Computer Science and Engineering, Bapatla Engineering College (Autonomous), Bapatla, India
e-mail: danielravuri@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_65

Keywords COVID-19 · SARA-CoV-2 · Exploratory data analysis (EDA) · Artificial intelligence (AI) and machine learning (ML)

1 Introduction

Since 24th April of 2020, 8,23,626 citizens with 40,598 confirmed deaths, and almost every person in the country is contaminated from SARA-CoV-2 infection [1]. Technological innovations are essential to the same battle against the epidemic; despite, there is a need to disrupt infection rates and the increased scope including its flu epidemic. ML and artificial intelligence (AI) interventions which include some 200 relevant papers including medical journals from 1st January to 24th April 2020 are already used in COVID-19 categories. Nevertheless, large-scale statistics, as well as prototype presenting, organization confirmation, and changes in local environments, are required for AI technologies dealing with COVID-19 to provide a massive impact. It further calls for international cooperation and equality, and indeed the participation of several other involved people, each with medical professionals.

Scientists and professionals are mostly from the regional and personal information sciences, even from Africa and Latin America, such as the African global network sciences, which should be active to maintain a positive effect on any AI program performed at the world scale. To do anything, technologies will not negotiate data confidentiality and security should be developed in front of the slow internet demand. For instance, medical diagnostic devices without any kind of integration. All these applications create the possibility to endorse cutting-edge AI virtualization as well as other environment-friendly challenges to the concerned use of quantitative services.

It is still essential to clarify the involvement of AI in presenting meaningful solutions to the identified epidemic. Interestingly, a worldwide research initiative must be constructed to initiate the measures against this pandemic but instead in the future epidemic without laying anybody behind it. One can fairly assume which in the comment on this thread-Coronavirus era, the world would become more interactive now than before, but also that AI has become more and more one of several drivers of modern civilization [2]. This virus outbreak highlights the urgent need to implement principles of AI effectively [3] by relevant parties. ‘Solidarity’ is indeed the motto of the world health organization’s international treatment program to guide select an acceptable cure against COVID-19 [4]. Confidence in each neighborhood inspires healthy distance, which is crucial to prevent the transmission of the infection within populations. Likewise, the presumption of solidarity could perhaps facilitate the development and promotion of imaginative and socially responsible AI applications in combating COVID-19 and now the worldwide environmental sustainability objectives [5].

AI systems from investigation towards control, some of the issues regarding AI systems is that it does not understand the problem or even how best to make a concerted effort. Further collaboration around researchers and the AI society are important. AI organization, which includes public servants, medical practitioners, and the first emergency workers, has already been seeking and thus should keep seeking support and assistance from consultants. Individuals could indeed encourage signal methodologies which are specialized software applicants. Recurrent or time-consuming activities encompass figuring designs in fabulously wealthy pictures, video, audio, or clinical research information. For example, tomography tests or activities requiring their aggregation of huge datasets from places such as diagnosis monitoring or professional network.

2 Literature Survey

An evaluation of new researchers at both the AI/COVID-19 interplay by several of us [6] provides a mechanism in which interdisciplinary investigation can be categorized on three factors: molecular, medical, and environmental (pharmacology and Infodemics). Biomedical implementations comprise simulation of protein structure [7], studies on viral nucleic acid [8], modifying medicine [9]. But instead, the revelation of medical products [10] most such developments use a variety of AI mechanisms together with specification and extraction of molecular biological medical databases, deep learning systems to forecast protein sequence attributes or protein–ligand binding inclinations, and its use of document learning models to genetics. Prediction of perspectives for genetic information including the use of clinical applications strengthened learning.

Health strategies to reduce treatment plans vary throughout the assessment to medical evaluation, rehabilitation, and assessment of the overall result. Deep learning frameworks would assist COVID-19 image-based diagnostics via trends in ultrasound images and imaging scans [11].

Natural-in-the-loop AI algorithms were established to reduce the waiting time for the examination of radiologists. Smartwatches, digital devices, numerous different wearable technologies, as well as other predictive maintenance software applications that enable physicians to validate patients from afar to save time, and safety precautions can effectively have been used to evaluate disease [12]. Each combination of data references which including medical history and medical imaging can help in assessing the consequences of patients [13] and could be utilized to show conventional healthcare necessities as with increased population for primary care unit rooms. Automation also might be useful for telecom and several other activities such as sanitizing and sterilization of surgical equipment.

AI can modify surveillance systems and observational data simulations besides epidemiology [14]. Specifically, AI is used to recognize and optimize treatment strategies in government health policies, such as preventing, social separation, and reconsideration and to enhance the traditional method of epidemiology by resembling components not widely recognized for diagnostic computations of transmission [15, 16]. Unattended clustering strategies and asset-scoring heuristics can further help determine similitude among both provinces and predict whether each location might need more resources, to incorporate multiple sources of information.

Besides, Infodemics and computer vision is being used to resolve misconceptions and deception through the management of the current knowledge overflow, causing outrage and making it very difficult to locate trustworthy sources [17, 18]. Analyzing audiences (e.g., media platforms, TV, radio) while optimizing assumption checks [19] could be helped with modeling techniques. For example, social media research provides some insights throughout environmental changes and distractions all over the infection and therefore its social; and cultural consequences [20]. The on-going work is to recognize the increase and dissemination of inciting hatred and defensive response on organizations and populations experiencing discrimination which could result in the actual act of violence in medicine [21]. Mostly in COVID-19 address, web portals and virtual agents could be able to propagate trustworthy feedback in scale [22] which needs frameworks for disseminating notifications to facts.

The performance appraisal method for the emerging global epidemic includes teamwork and priority [23], which focuses on either the key unsatisfied conditions that consider organizational reality. Some of the approaches must be feasible and focused on implementing processes [24] most of which are proven and must resist irreversibly damaging the patient care with innovative technologies that might not necessarily lead to improving efficiency. Moreover, prospective and current reforms are needed to meet the particular needs and circumstances of regions of the world with different levels of economic development. Patient-centered care service offerings ought not to neglect effective capitalist structures to maintain proper, health and safety minimizing risk, and damage potential. Priority given to combat COVID-19 is thorough the evaluation of alternative solutions remains mandated, which could be quickly pursued against risking.

To stimulate innovation, manufacturing, and high availability of resources for COVID-19 [25] progress in a way to achieve the advanced health based on information technology. Cooperation with the public health players, government sector participants, and other collaborators are also underway. Collaboration and cooperation among both government agencies and multilateral organizations will indeed facilitate the channeling of ground research that would help introduce technologies in relatively weak economic policy and institutional arrangements countries. Comparing the integrity, therapeutic reliability, and production adequacy of healthcare-related application areas using AI would allow for formulating an effective process of making decisions yet would procedure permits with guidelines [26, 27].

Combating COVID-19 worldwide Infodemics as a scientific concept should be viewed at the same time as the outbreak actually, as the shift in behavior is important to just the global epidemic reaction. News and mass communication stay significant, and health insurance premiums of deception and distortion transmission have to be quantified. Infodemic detection strategies could also be used to facilitate the immediate traditions of facts to the local community, communication, and context-specific information and intervention. AI methods and tools for filling individuals' and politicians' health information gaps should be used for a society-wide response based on evidence and science. The regional and global community should therefore convey and broaden professional exercise, establish guidelines, enabling collaborations, and focus on providing guidance and technical help to the authorities and the appropriate national policymakers to maintain international peace and security to address Infodemics productively while guaranteeing its essential right to communicate.

The global epidemic promotes xenophobia, hatred, and discrimination, raising a pervasive, and likely long-term social justice threat [28]. Changing the nature of violence and promoting, it will help to develop more successful solutions and change the result.

3 Methodology

3.1 Communicating Medical Information as Well as a Framework

Investigations are necessary for smart devices, dozens of data-sharing projects throughout COVID_19, covering the international, national, and local thresholds, are occurring right mostly in three dimensions. Such metrics usually involve genetic sequence [29], genomic analyses [30]. Protein compositions, medical evidence for patients, and medical imaging details for the occurrences of pharmacological information [31]. The ultra-fragmentation of intelligent exchange of data efforts is a major obstacle because it may lead to progress that would be confined to individual programs and entire communities. The improvement and distribution of modern innovations might be accelerated by traceability interventions for statistics, template, and reliability requirements. During this point, international data retention strategies, which are free, inclusive, and compatible, and checked, will repair damaged and foster collaboration among different inhabitants of the area geographies [32].

Fully accessible scientific knowledge will speed up the proliferation of awareness through intersectional AI collaborations around state lines and service delivery of nationwide health organizations. Timely identification, authorization, and impact analysis on nutrition allow possible hazards incorporating open source information (for instance, Epidemic Intelligence Open Source (EIOS) [33] network's location information application). This same social network for health protection recognizes governments, international agencies, and academic institutions which either interdependently evaluates or accepts precise data on epidemics events under that same ideology of cooperation and not rivalry for automatic recognition. The EIOS service published that the very first post on 31st December 2019, documenting an unexplained outbreak of respiratory infections in Wuhan.

Common services and interconnectivity among both repositories can allow for organized intervention and decision-making either at regional, national, and municipal stages of development mostly from the epidemiological aspect. Acknowledging the epidemiological mechanisms and vulnerability characteristics of various influenza populations, while focusing on the multiple moments, may include consideration of national health capability, public policy strategies, ecosystem processes, and social impacts of COVID-19.

Limitations must be conquered, due to the special; technical, architectural conditions, absence of records; characteristics of authentication and explanation; security concerns about personal information and confidential information, and connectivity needs. The collaboration of word embedding and proven predictive analytics might also speed up approaches' adaptive response to different societies. Examples of commonly accessible simulations comprise visual analysis models, prediction of patients' outcomes, filtering misconceptions, and intelligence based on variations of dissemination across virtual media or distilling expertise statistics by broad research publications. Artificial intelligence performance management software packages that reflect the reality of acceptable, interpersonal, therapeutic, law are needed to achieve real-world for open predictive analytics.

3.2 Digital Collaboration and Regional Priorities

Additionally, advanced analytics for tackling age-old issues were introduced. However, it does not need to denote developers should have the possibility to build such application areas. Any AI framework for countering COVID-19 should be checked to make sure it follows ethical standards and values human dignity in particular. Besides, specialists are constantly faced with suggestions about protecting basic rights, including the right to data protection, even in compliance with national security and rationality-by-design standards to create, and deploy AI-driven remedies. When implementing either of those advanced analytics on a global level, there is a need to certify that they will always not infringe violate global democratic rights obligations, which include anti-discrimination responsibilities, surveillance prevention, and editorial sources protection. Decision-makers might also help ensure that AI-friendly approaches are central to beliefs such as diversity and affordability. AI medicare beneficiaries will also encourage equal for balanced access to the global healthcare service system to support participating countries' national health obligations.

The complex nature of both the epidemic calls for systemic approaches, but organizational change is also important to consider biases and contexts. In locations with lower system implementation of chronic disorders with relevant picture configurations, including infectious diseases like tuberculosis and HIV, a system besides diagnosing COVID-19 pneumonia will have to be radically different training. Similarly, separate socioeconomic, cultural, and contextual possibilities than some outlined by academic literature—mainly developed for China or western countries—should indeed be put into consideration by statistical measures customized to developing countries, island countries, humanitarian assistance, or unstable countries. Chatbot also needs analytical patterns for language acquisition to advise primary prevention and sometimes large volumes of education information available for very few hundred from out 7000 linguistic groups today.

Table 1 Data collection based on the attributes

Data	Attributes
Age data	Age group, total cases, percentage
Hospital data	State, NumPrimaryHealthCenters_HMIS, NumCommunityHealthCenters_HMIS, NumSubDistrictHospitals_HMIS, NumDistrictHospitals_HMIS, TotalPublicHealthFacilities_HMIS, NumPublicBeds_HMIS, NumRuralHospitals_NHP18, NumRuralBeds_NHP18, NumUrbanHospitals_NHP18, NumUrbanBeds_NHP18
Testing lab	Lab, address, Pincode, city, state, type
Covid data	Date,Time,State/UnionTerritory, ConfirmedIndianNational, ConfirmedForeignNational, Cured, Deaths, Confirmed
World data	Province/State, Country/Region, Lat, Long, Date, Confirmed, Deaths, Recovered, Active, WHO Region
Making predictions (India)	Province/State, Country/Region, Lat, Long, Date, Confirmed, Deaths, Recovered, Active, WHO Region

4 Result Analysis

Initially load the data using the Kaggle dataset, then normalize data, statewide index the data using set_index, attributes are listed to analyze the data, handles the missing data. The following are the tables that show the considered data with considered attributes for COVID-19 worldwide and India using the prophet model for better predictions. For exploratory data analysis, this paper performs forecasting of data using prophet model libraries. The use of the Prophet.make_future_dataframe will extend the data frame depending on the days specified.

Table 1 describes the information of obtained COVID-19 attribute data for future predictions Table 2 describes the information of obtained COVID-19 Indian attribute data for future predictions using the prophet model. Table 3 describes the information of obtained COVID-19 Worldwide attribute data for future predictions using the prophet model. Figure 1 describes the plot diagram for overall active confirmed, recovered, and death cases in India. Figure 2 describes the plot diagram for predicted active confirmed, recovered, and death cases in India.

5 Conclusion

Several of the proposals and services proposed have still been appropriately designed to either be operational, despite the multiple implementations that offer different development, testing, and delivery probabilities. Consequently, for patients and the automation society, it is necessary to recognize the increasing advancements, which

Table 2 Indian COVID-19 data using a Prediction model

COVID-19 India	Attributes	Model
Confirmed Cases in India	Date,Confirmed	model = Prophet()
Making predictions	ds(date), trend, yhat_lower, yhat_upper, trend_lower, trend_upper, additive_terms, additive_terms_lower, additive_terms_upper, weekly,weekly_lower, weekly_upper, multiplicative_terms, multiplicative_terms_lower, multiplicative_terms_upper, yhat	forecast_india_conf = model.predict(future)
Recovered cases in India	ds(date), trend, yhat_lower, yhat_upper, trend_lower, trend_upper, additive_terms, additive_terms_lower, additive_terms_upper, weekly, weekly_lower, weekly_upper, multiplicative_terms, multiplicative_terms_lower, multiplicative_terms_upper, yhat	forecast_india_recover = model.predict(future)
Deaths in India	ds(date), trend, yhat_lower, yhat_upper, trend_lower, trend_upper, additive_terms, additive_terms_lower, additive_terms_upper, weekly, weekly_lower, weekly_upper, multiplicative_terms, multiplicative_terms_lower, multiplicative_terms_upper, yhat	forecast_india_death = model.predict(future)

will help to react shortly, improve mid-term, and plan for potential infectious diseases by implementing artificial intelligence predictions. This article describes the artificial intelligence requirements as well as the categorization of evolutionary COVID-19 data, artificial intelligence-based communication for medical information, control, and digital collaboration by concerning the regional priorities. Results for COVID-19 were tested and predicted by using exploratory data analysis and prediction analysis through a global database of COVID-19 medical data.

Table 3 Worldwide COVID-19 data using a prediction model

COVID-19 World	Attributes	Model
Making predictions (world)	Province/State, country/Region, lat, long, date, confirmed, deaths, recovered, active, WHO region	model = Prophet()
Confirmed cases in world	ds(date), trend, yhat_lower, yhat_upper, trend_lower, trend_upper, additive_terms, additive_terms_lower, additive_terms_upper, weekly, weekly_lower, weekly_upper, multiplicative_terms, multiplicative_terms_lower, multiplicative_terms_upper, yhat	forecast_world_conf = model.predict(future)
Recovered cases in world	ds(date), trend, yhat_lower, yhat_upper, trend_lower, trend_upper, additive_terms, additive_terms_lower, additive_terms_upper, weekly, weekly_lower, weekly_upper, multiplicative_terms, multiplicative_terms_lower, multiplicative_terms_upper, yhat	forecast_world_recover = model.predict(future)
Deaths in world	ds(date), trend, yhat_lower, yhat_upper, trend_lower, trend_upper, additive_terms, additive_terms_lower, additive_terms_upper, weekly, weekly_lower, weekly_upper, multiplicative_terms, multiplicative_terms_lower, multiplicative_terms_upper, yhat	forecast_world_death = model.predict(future)

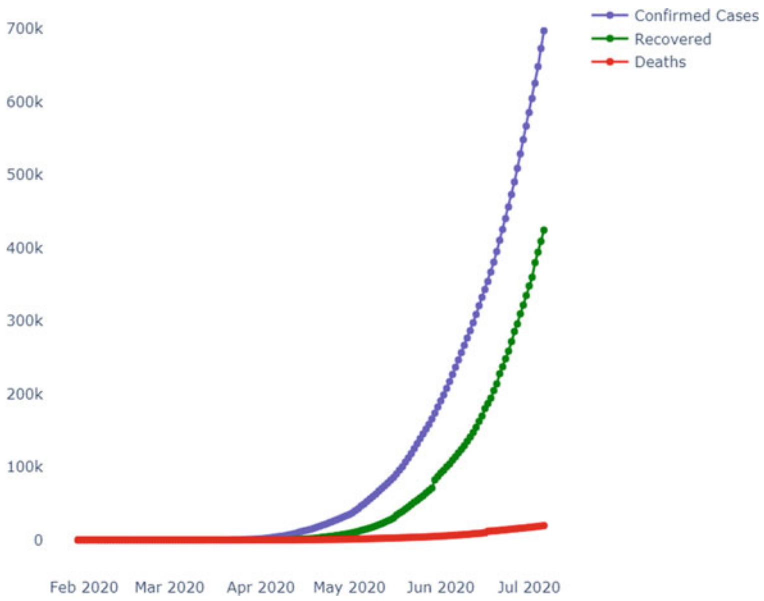


Fig. 1 Plot for total cases in India

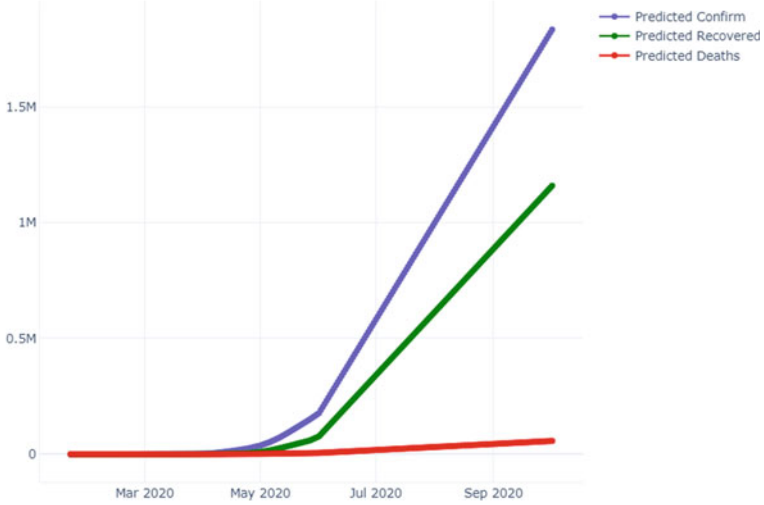


Fig. 2 Plot for predicted cases in India

References

1. WHO (2020) Coronavirus disease (COVID-2019) situation reports. https://www.who.int/emergencies/disease/novel-coronavirus-2019/situation-reports
2. UN Secretary General's High-level Panel on Digital Cooperation (2019) The age of digital interdependence
3. Jobin A, Lenca M, Vayena E (2019) Nat Mach Intell 1:389–399
4. WHO (2020) Solidarity" clinical trial for COVID-19 treatments. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov/solidarity-clinical-trial-for-covid-19-treatments
5. Vinuesa R (2020) Nat Commun 11:233
6. Bullock J, Luccioni A, Hoffmann Pham K, Lam CS, Luengo-Oroz M (2020) Preprint at https://arxiv.org/abs/2003.11336
7. Senior AW et al (2020) Nature 577:706–710
8. Lopez-Rincon A et al (2020). Preprint at. https://doi.org/10.1101/2020.03.13.990242
9. Richardson P et al (2020) Lancet 395:E30–E31
10. Zhavoronkov A et al (2020) Preprint at https://doi.org/https://doi.org/10.26434/chemrxiv.11829102.v2
11. Shi F et al (2020). IEEE Rev Biomed Eng. https://doi.org/10.1109/RBME.2020.2987975
12. Radin J, Wineinger NE, Topol EJ, Steinhubl SR (2020) Lancet dig. Health 2:E85–E93
13. Shi W, Peng X (2020). Preprint at. https://doi.org/10.2139/ssrn.3546089
14. AI-qanessMAA, Ewees AA, Fan H, AbdEI Aziz MJ (2020) Clin Med 9:674
15. Nuria O et al (2020). Sci Adv. https://doi.org/10.1126/sciadv.abc0764
16. Dandekar R, Barbastathis G (2020) Preprint at https://arxiv.org/abs/2003.09403
17. WHO (2020) Infodemic management: a key component of the COVID-19 global response. https://apps.who.int/iris/handle/10665/331775 (2020)
18. WHO (2020) Coronavirus Disease 2019 (COVID-19)—Situation Report 86
19. WHO (2020) Coronavirus Disease 2019 (COVID-19)—Situation Report 100
20. Gallotti R et al (2020). Preprint at. https://doi.org/10.1101/2020.04.08.20057968
21. Velasquez N et al (2020) Preprint at https://arxiv.org/abs/2004.00673
22. Sundareswaran V, Firth-Butterfield K (2020) Chatbots provide millions with COVID-19 information everyday, but they can be improved- here's how. World Economic Forum https://www.weforum.org/agenda/2020/04/chatbots-covid-19-governance-improved-here-s-how/
23. WHO (2020) Digital health. https://www.who.int/health-topics/digital-health
24. Topol EJ (2019) Nat Med 25:44–56
25. WHO (2020) Access to COVID-19 Tools (ACT) Accelerator
26. The CONSORT-AI and SPIRIT-AI Steering Group (2019) Nat Med 25:1467–1468
27. Collins GS, Moons KGM (2019) Lancet 393:1577–1579
28. UN Human Rights (2020) Press briefing note on Americas/prison conditions. https://shar.es/aHIzht
29. Cohen J (2020) Science. https://www.sciencemag.org/news/2020/01/chinese-researchers-reveal-draft-genome-virus-implicated-wuhan-pneumonia-outbreak
30. Hadfield J et al (2018) Bioinformatics 34:4121–4123
31. Zastrow M (2020). Nature. https://doi.org/10.1038/d41586-020-00740-y
32. Teran J (2020) UN Department of Economic and Social Affairs. https://covid-19-response.unstatshub.org/open-data/hdx-making-covid19-data-accessible/
33. WHO (2020) Epidemic intelligence from open sources. https://www.who.int/eios
34. GitHub (2020) UNGlobalPulse/covid19-literature-search. https://github.com/UNGlobalPulse/covid-19-literature-search
35. Wang LL et al (2020) Preprint at https://arxiv.org/abs/2004.10706

Design VLSI Architecture for 2_D DWT Using NEDA and KSA Technique

Satyendra Tripathi, Bharat Mishra, and Ashutosh Kumar Singh

Abstract Image compression emanating utilizations of data compression on digital images. The Discrete Wavelet Transform (DWT) is getting familiar with an extreme move in image processing. To eliminate the drawback in the JPEG standard and rising areas of mobile and Internet communications, the new JPEG2000 standard has been created based on the principles of DWT. With the prevalence of the 2_D DWT, the technology has seen predictable improvement. All things considered, there is still a need for more robust and efficient compression technology using DWT. Hence, a 2_D DWT is implementing with a multiplier-less NEDA technique. NEDA is consisting of a buffer, ROM, and adder. There are many types of adders used in digit circuits but Kogge Stone Adder (KSA) is used in the proposed method. KSA is a very efficient adder and consists of the XOR gate and HA. Proposed Schemes are simulated Xilinx software with the VHSIC Hardware Description Language (VHDL) platform and calculates parameter i.e. speed and frequency.

Keywords 2_D discrete wavelet transform (DWT) · New efficient distributive arithmetic (NEDA) · Read only memory (ROM) · Half adder (HA) · Kogge stone adder (KSA)

1 Introduction

A portion of its highlights are; versatile time-recurrence windows, lower associating mutilation for signal handling [1, 2]. As a result of these highlights 1_D DWT [3, 4] and 2_D DWT [5] can be utilized in different applications, for example, numerical examination [6], analysis of various signals, image coding, pattern recognition, watermarking. It can also be used in Biomedicine [7]. Lots of calculations and plans have been prompted during the most recent three decades to make an appropriate

---

S. Tripathi (✉) · B. Mishra
MGCGV Chitrakoot, Chitrakoot, India
e-mail: satyendra.mgcgv16@gmail.com

A. K. Singh
IIIT-Allahabad, Prayagraj, Allahabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4_66

equipment execution of 1_D DWT [8] and 2_D DWT. DWT gives a proper computing technique for the representation of a wide sample of signals. DWT gives sufficient information about the synthesis and testing of the original signal with a reduction in time.

The important guideline is the end of monotonous signs and information that can reduce the methodology. From a logical point of view, adjust the pixel framework of 2-D [9] size, into a scientific, numerically uncorrelated informational index. The change in the information can be seen previously, and then the picture is transmitted. Later when the picture will be needed, it is straightforwardly decompressed to revamp the ordinal one. There may be a slight difference of information from the unique picture; however, the measure of data is the same. The attention on picture pressure was in presence for years now. At first, the examination takes a shot at the picture pressure method that was centered on diminishing the span of video for broadcasting. The utilization of picture pressure in such a situation, won't diminish the measure of the mixed media record, and also exceeds the speed of transmission to get successful and productive usage of transmission of data.

In any case, locale and speed are large conflicting necessities to improve speed comes to fruition generally in greater areas. Earlier, control usage was helping conversely to improve the area and speed. Regardless, control is being given more noteworthiness as zone and speed because of sublime advancement of minimized and remote handheld blended-media contraptions. The usage of power is the most essential factor for these devices [10].

2-D DWT is used to many applications like medical image, tomography, speech and voice, pattern recognition, biomedicine, and computer graphics. 2-D DWT is a technique that divided the information in the multiresolution frequency domain. 2-D DWT is divided information in four parts i.e. normal, normal distinction, contrast normal and contrast distinction is present Fig. 1. G means the average of two samples and H means the difference between two samples.

2 Literature Review

Samit et al. [1], abridge the circuit intricacy and enhance the execution are proposed; DWT is applied to the periodic and non-periodic signal that used to speech and video signal. This article designed 9/7 DWT based on DA and KSA. The drawback of this paper is KSA does not properly work in all the binary bits and sources of info, which profoundly expands the circuit unpredictability. Proposed procedures depend on rapid region productive 2-D DWT utilizing 9/7 channel based altered MDA technique also, KSA. MDA strategy is connected to the bottommost and eminent pass channel of the DWT. This method comprises a viper, move to enroll, and handout of a multiplier. Plan and result are actualized in Xilinx programming and confirmed resistor exchange level and waveform.

Biswas et al. [2], proposed a programmable one-dimensional discrete wavelet bundle change processor. Contrasted and existing designs, the proposed processor can

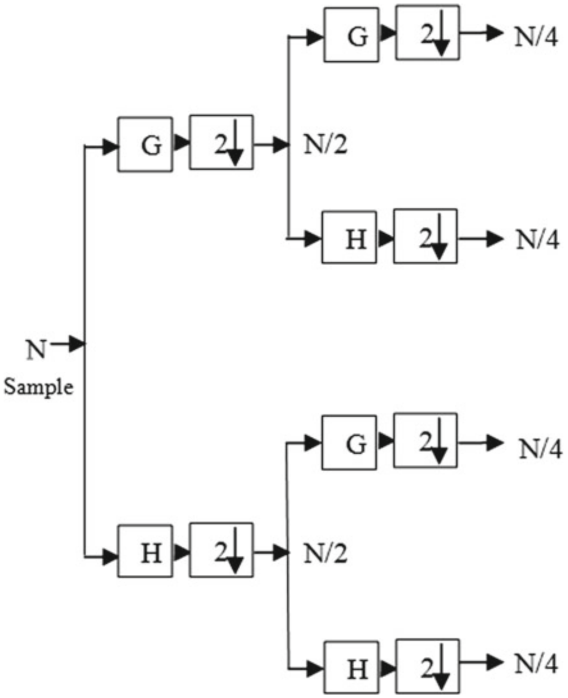


Fig. 1 2-D DWT

complete both wavelet bits of information to be deserted are found. In any case, the picture pressure innovation has turned into a mainstream decision for dealing with the very mind-boggling spatial determination of the current picture catching sensors. The fantastic video transmission for TV and media has additionally lifted the utilization of picture pressure. In addition, picture pressure attempts an important association in various basic and incalculable applications, for example, video conferencing, TV, biomedical imaging, remote detecting (utilizes satellite symbolism for atmosphere and earthly data total), FAX, space investigation, records handling, and transmission, controlling remote vehicle for military, squander administration, and so on.

Mamatha et al. [6], the previously mentioned application particularly demonstrates the necessity for sufficient capacity, significant data transfer capacity for transmission, and delayed transmission period for a picture, sound, and video signals. In this present circumstance, the innovation that can be gotten to and executed under these conditions is the picture pressure. Picture pressure gives an answer for every one of the difficulties looked in the above zones. It lessens the measure of sight and sound information required for capacity and transmission. It packs the information, at that point stores it, or transmits it. At the point, when the information is required to be perused, at that point the picture decompression replicates the first data. For example, when the pressure proportion is 32:1, at that point the capacity zone, exchange speed, data transfer capacity utilization, and transmission postpone requirements can be limited by a factor of 32, with no misfortune in quality. The portrayal of the photo is first changed from the given spatial space to a substitute area with the assistance of well-known flag changing and coding procedures.

Martina et al. [7], the existing method consumed more memory is used to storing the intermediate computational multiplier based DWT. The DWT needs to process gigantic measures of information at high speed, making it unsuitable for real-time image compression with better performance. The objective of the research work is to infer effective designs for the equipment execution of the multiplier-less DWT, compared to the existing architectures.

Kumar et al. [11], the optimization factors are the speed, low area, and hardware multifaceted nature required, with the size of the given information picture and the necessary degrees of deterioration. The MDA based DWT design for the 9/7 wavelet channel coefficient multiplier-less engineering is enhanced, to accomplish better speed and higher equipment usage, by utilizing a solitary clock to focus on tasks updates. In the cutting edge mechanical industry, it is very difficult to go over a field that isn't impacted by the advanced picture preparing. It assumes a vital part of all specialized areas, somehow. Keeping in mind the end goal to keep up a conservative portrayal, just a few applications are depicted that are identified with the proposed work. By and large, the fields that make utilization of advanced picture handling can be divided into morphology, criminology, photography, microscopy, biomedical imaging, remote detecting, transportation, explore sciences, military application, and numerous others.

3 Proposed Methodology

Multiplier-less NEDA and KSA technique are shown in Fig. 2. In this figure, input morsel is passed through shift registers and all input–output is added symmetrically with the help of adder. The LPS and HPS input are passed through NEDA and KSA technique and the output becomes Y_{HPS} and Y_{LPS}.

3.1 Example of Multiplier-Less HPS

If it takes the HPS coefficients such as g_0, g_1, $g_{2,}$ and g_3 multiply by v_1, v_2, $v_{3,}$ and v_4 then multiplier-less 1-D DWT [12, 13] HPS output is

$$Y_{\text{HPS}} = \begin{bmatrix} g_0 & g_1 & g_2 & g_3 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$$

where,

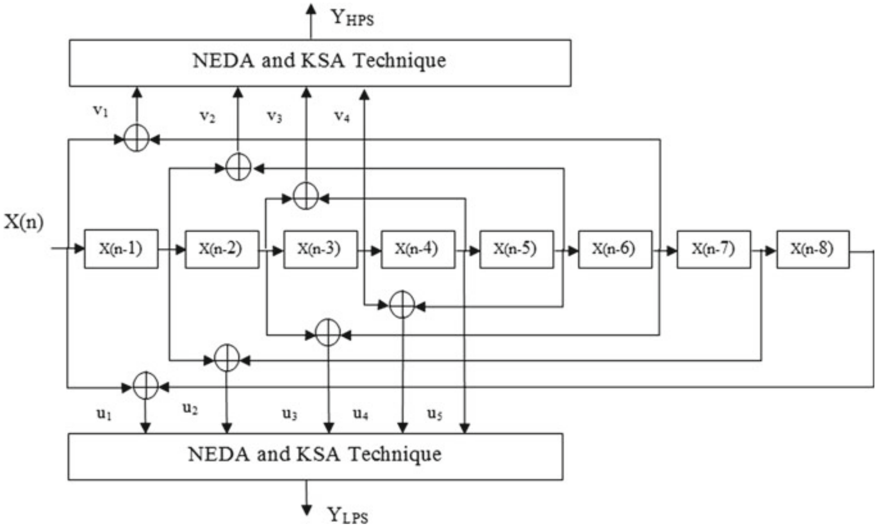


Fig. 2 Proposed 1-D DWT using NEDA and KSA technique

$v_1 = X(n) + X(n-6)$

$v_2 = X(n-1) + X(n-5)$

$v_3 = X(n-2) + X(n-4)$

$v_4 = X(n-3).$

$$Y_{\mathrm{LPS}} = \begin{bmatrix} 71 & -38 & -4 & 6 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$$

So,

$$Y_{\mathrm{HPS}} = \begin{bmatrix} 0\,1000111 & 11011010 & 11111100 & 00000110 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$$

All the HPS coefficient arranges down to up is below:

$$Y_{\mathrm{H}} = \begin{bmatrix} 1\,0\,0\,0 \\ 1\,1\,0\,1 \\ 1\,0\,1\,1 \\ 0\,1\,1\,0 \\ 0\,1\,1\,0 \\ 0\,0\,1\,0 \\ 1\,1\,1\,0 \\ 0\,1\,1\,0 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$$

All rows pass through look up table and replace the LPS coefficient to input

$$Y_{\mathrm{H}} = \begin{bmatrix} 1\,0\,0\,0 \\ 1\,1\,0\,1 \\ 1\,0\,1\,1 \\ 0\,1\,1\,0 \\ 0\,1\,1\,0 \\ 0\,0\,1\,0 \\ 1\,1\,1\,0 \\ 0\,1\,1\,0 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_1 + v_2 + v_4 \\ v_1 + v_3 + v_4 \\ v_2 + v_3 \\ v_2 + v_3 \\ v_3 \\ v_1 + v_2 + v_3 \\ v_2 + v_3 \end{bmatrix}$$

Let $v_1 = 4$, $v_2 = 4$, $v_3 = 4$ and $v_4 = 2$ and put above equation and last row value is represented by 2's complement value.

Then

$$\begin{aligned} &KK_1 = v_1 = 0100,\ \ KK_2 = v_2 + v_2 + v_4 = 1010,\ \ KK_3 = v_1 + v_3 + v_4 = 1010, \\ &KK_4 = v_2 + v_3 = 1000,\ \ KK_5 = v_2 + v_3 = 1000,\ \ KK_6 = v_3 = 1000 \\ &KK_7 = v_1 + v_2 + v_3 = 1100,\ \ KK_8 = v_2 + v_3 = \text{not }(1000) + \ ''0001'' = 1000 \end{aligned}$$

All KK_1 to KK_8 value passed through sign extension block then

$$\begin{aligned} &KK_1 = v_1 = 00100, KK_2 = v_2 + v_2 + v_4 = 01010, KK_3 = v_1 + v_3 + v_4 = 01010, \\ &KK_4 = v_2 + v_3 = 01000, KK_5 = v_2 + v_3 = 01000, KK_6 = v_3 = 001000, \\ &KK_7 = v_1 + v_2 + v_3 = 01100, KK_8 = v_2 + v_3 = not(01000) + \ ''0001'' = 11000 \end{aligned}$$

KK_1 is left shift one bit and add KK_2 and store output YY_1

$$\begin{aligned} &= 0'00100 \\ &\ \ + 01010 \\ YY_1 &= 011000 \end{aligned}$$

YY_1 is left shift one bit and add KK_3 and store output YY_2

$$\begin{aligned} &= 0'011000 \\ &+ 01010 \\ YY_2 &= 1000000 \end{aligned}$$

YY_2 is left shift one bit and add KK_4 and store output YY_3

$$\begin{aligned} &= 0'1000000 \\ &+ 01000 \\ YY_3 &= 10000000 \end{aligned}$$

YY_3 is left shift one bit and add KK_5 and store output YY_4

$$\begin{aligned} &= 0'10000000 \\ &+ 01000 \\ YY_4 &= 100000000 \end{aligned}$$

YY_4 is left shift one bit and add KK_6 and store output YY_5

$$\begin{aligned} &= 0'100000000 \\ &+ 00100 \\ YY_5 &= 0110000000 \end{aligned}$$

YY_5 is left shift one bit and add K_7 and store output YY_6

$$\begin{aligned} &= 0'11100000000 \\ &+ 01100 \\ YY_6 &= 100100000000 \end{aligned}$$

YY_6 is left shift one bit and add KK_8 and store output YY_7

$$\begin{aligned} &= 0'10010000000 \\ &+ 11000 \\ Y_7 &= 1000010000000 \end{aligned}$$

Final output $Y_{\text{HPS}} = YY_7 = (0000010000000)_2$ (Carry Reject).

4 Simulation Result

Figure 3 is the view of the technology of first level DWT. Here ‘e’ is the input of 4-bit, clk is the clock signal applied and here two output Comes each of 12-bit.one is ‘yh’ means the output for the eminent pass filter and the other is; yl’ means the output for the bottommost pass filter.

Figure 4 shows the RTL view of first level DWT. This view includes the shift registers, KSA, D-flip flops, and all its components. First, the input passes through D-flip flops, and then symmetrically addition is performed through KSA and then after using the NEDA technique, the final output comes.

Figure 5 shows the view of the technology of second level DWT. Here the input is of 4-bit and the two outputs are of 19 bit. One is ‘yh1’ which is the output for the eminent pass filter and the other is ‘yl1’ which is the output for the bottommost pass filter.

Figure 6 shows the RTL view of 2-D DWT. It has all the components of 2-D DWT. It contains all the shift registers, D-flip flops, BK adder [14]. This RTL schematic depends on the view of technology.

Figure 7 shows the waveform of the 2_D DWT. Here the input is given as ‘0011’ and the output finally comes for both the filters. ‘yh’ is ‘1111101000000000000’ for eminent pass filter output and ‘yl’ is ‘0000011000000000000’ for the bottommost pass filter output.

In this Table 1, when it comes to 1_D and 2_D DWT in the case of ‘a number of slices’ the proposed delineation is 56.6% exceptional than the previous delineation, in the case of ‘number of flip flops’ the proposed delineation is 33.3% exceptional than the previous delineation, in case of ‘number of LUTs’ the proposed delineation is 29.2% exceptional than the previous delineation, in case of ‘maximum delay’ the proposed delineation is 25.19% exceptional than the proposed delineation. Similarly,

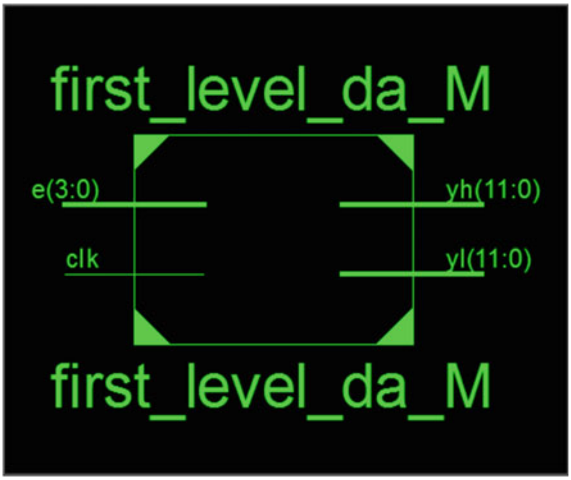


Fig. 3 VTS for 1-D DWT

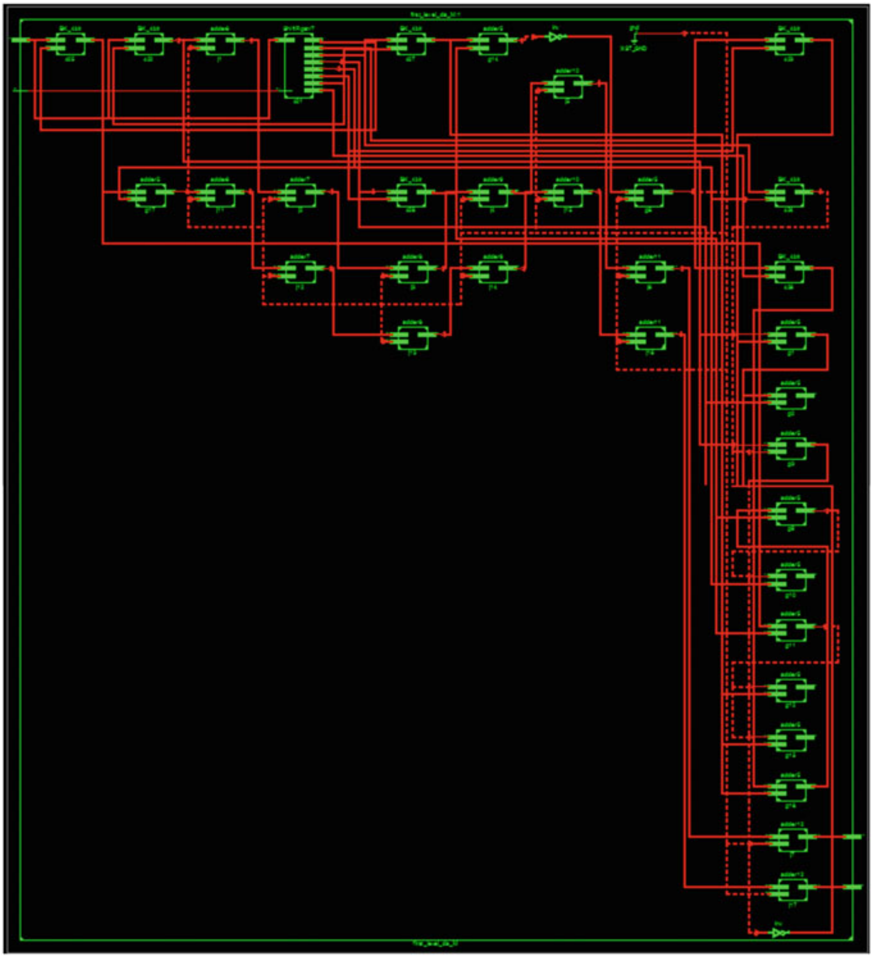


Fig. 4 RTL for 1-D DWT

when it comes to 2-D DWT, the proposed delineation is much better than the previous delineation.

5 Conclusion

The Discrete Wavelet Transform is being designed using the New Efficient Distributive Arithmetic technique and Kogge Stone Adder. This transform breaks the signal into a mutually orthogonal set of wavelets. It is the implementation of a wavelet transform in which it uses some discrete set of wavelets and translations after obeying

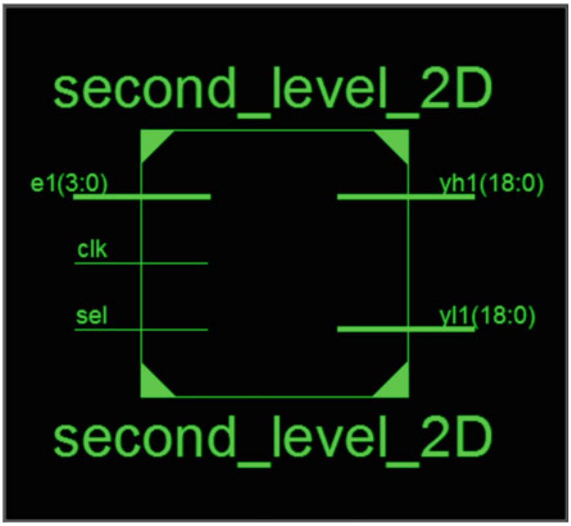


Fig. 5 VTS for 2-D DWT

some rules. The DWT provides complete information for analysis and synthesis of signals and also reduces the time. When it comes to 1_D and 2_D DWT in the case of ‘a number of slices’ the proposed delineation is 56.6% exceptional than the previous delineation, in the case of ‘number of flip flops’ the proposed delineation is 33.3% exceptional than the previous delineation, in case of ‘number of LUTs’ the proposed delineation is 29.2% exceptional than the previous delineation, in case of ‘maximum delay’ the proposed delineation is 25.19% exceptional than the previous delineation.

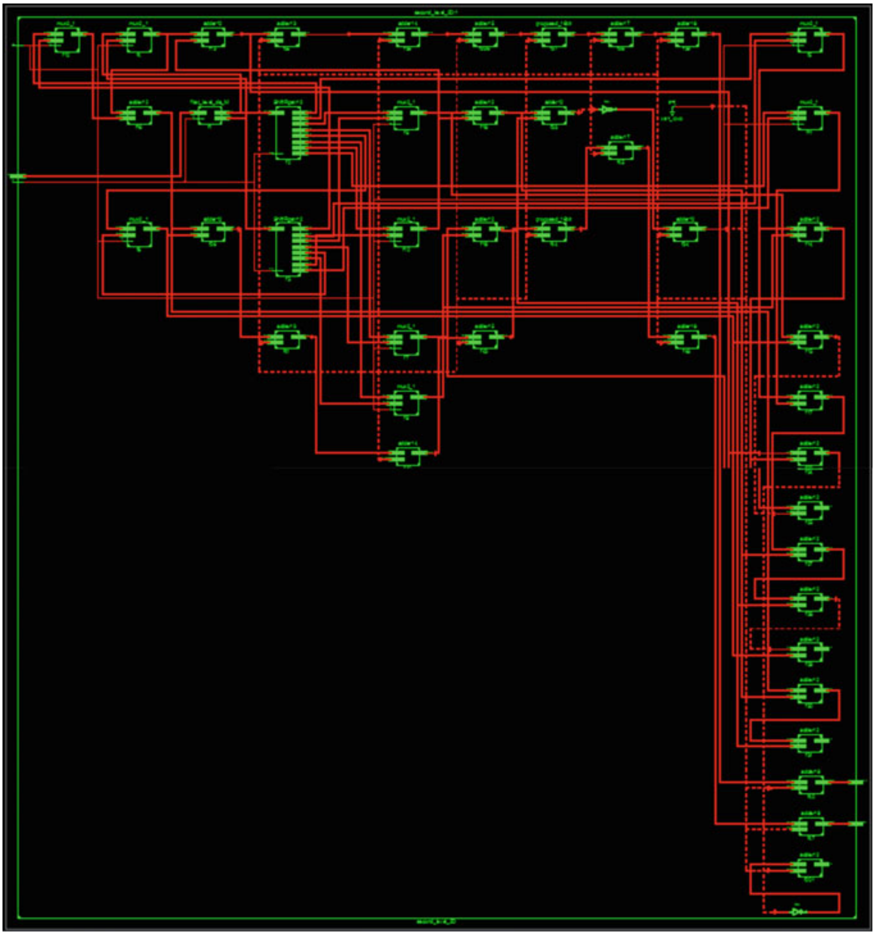
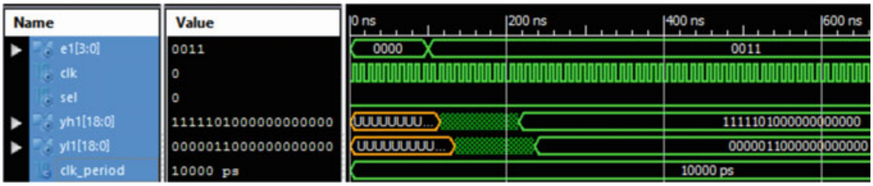


Fig. 6 RTL for 2-D DWT



Name	Value
e1[3:0]	0011
clk	0
sel	0
yh1[18:0]	1111101000000000000
yl1[18:0]	0000011000000000000
clk_period	10000 ps

Fig. 7 VHDL Test-bench in 2-D DWT

Table 1 Comparison of result with previous 2-D DWT implementation

Parameter	1-D DWT		2-D DWT	
	Previous design	Proposed design	Previous design	Proposed design
Number of slice	346	150	752	514
Number of slice flip flop	48	32	454	224
Number of LUTs	359	254	1393	899
MCPD	23.146 ns	17.316 ns	28.998 ns	

References

1. Dubey SK, Kourav AK, Sharma S (2017) High speed 2-D discrete wavelet transform using distributed arithmetic and Kogge Stone adder technique. In: International conference on communication and signal processing, April 6–8, India
2. Biswas R, Malreddy SR, Banerjee S (2017) A high precision-low area unified architecture for lossy and lossless 3D multi-level discrete wavelet transform. IEEE Trans Circ Syst Video Technol 45(5):1–11
3. Cao X, Xie Q, Peng C, Wang Q, Yu D (2006) An efficient VLSI implementation of distributed architecture for DWT. In: Proceedings of IEEE workshop on multimedia and signal process, pp 364–367
4. Tewari G, Sardar S, Babu KA (2011)High-Speed & Memory Efficient 2-D DWT on Xilinx Spartan3A DSP using scalable polyphase structure with DA for JPEG2000 Standard.978-1-4244-8679-3/11/$26.00 ©2011 IEEE
5. Martina M, Masera G (2006) Low-complexity, efficient 9/7 wavelet filters VLSI implementation. IEEE Trans Circuits Syst II, Expr Brief 53(11):1289–1293
6. Mamatha I, Tripathi S, Sudarshan TSB (2016) Pipelined architecture for filter bank based 1-D DWT. In: International conference on signal processing and integrated networks (SPIN), pp 47–52
7. Martina M, Masera G, Ruo Roch M, Piccinini G (2015) Result-biased distributed-arithmetic-based filter architectures for approximately computing the DWT. In: IEEE transactions on circuits and systems—I: regular papers, vol 62, No 8
8. Mohanty BK, Meher PK (2009) Efficient multiplierless designs for 1-D DWT using 9/7 filters based on distributed arithmetic. ISIC 2009
9. Mohanty BK, Meher PK (2011) Memory efficient modular VLSI architecture for high-throughput and low-latency implementation of multilevel lifting 2-D DWT. IEEE Trans Signal Proces 59(5)
10. Alam M, Rahman CA, Jullian G (2003) Efficient distributed arithmetic based DWT architectures for multimedia applications. In: Proceedings of IEEE workshop on SoC for real-time applications, pp 333–336
11. Mohanty BK, Meher PK (2013) Memory-efficient high-speed convolution-based generic structure for multilevel 2-D DWT. IEEE Trans Circ Syst Video Technol 23(2):353–363
12. Zhao X, Vi Y, Erdogan AT, Arslan T (2000) A high-efficiency reconfigurable 2-D discrete wavelet transform engine for JPEG 2000 implementation on next generation digital cameras. 978-1-4244-6683-2/10/$26.00 © 2010 IEEE
13. Baviskar A, Ashtekar S, Chintawar A, Baviskar J, Mulla A (2014) Performance analysis of sub-band replacement DWT based image compression technique. 978-1-4799-5364-6/14/$31.00 © 2014 IEEE
14. Ramkumar B, Kittur HM (2012) Low-power and area-efficient carry select adder. IEEE Trans Very Large Scale Integr (VLSI) Syst 20(2)
15. Sahoo R, Roy S, Chaudhuri SS (2014) Haar wavelet transform image compression using run length encoding. In: International conference on communication and signal processing, April 3–5, 2014, India

Author Index

A
Aashritha, D., 239
Adhikari, Bikram, 201
Akram, Faiz, 1
Akter, Rupa, 263, 279
Alabkari, Mariam, 363
Alam, M. J., 263, 279, 301
Al-azani, Samah Ali, 733
Al-Bahrani, Mustafa, 349
Albur, Wafa, 87
Al-Faraj, Saleh, 349
Al-Ghamdi, Saeed, 349
Alghamdi, Shahad, 363
Aljishi, Fatima, 363
Alsedrah, Shaikha, 87
Alufaisan, Shurouq, 87
Álvarez, Elia N. Cabrera, 77
Anitha, R., 601
Anonto, Rapsan Amin, 279
Anu Priya, P., 407
Arunnehru, J., 515

B
Baalamurugan, K. M., 753
Baligar, J. S., 647
Baranidharan, V., 501
Barua, Arnob, 301
Bashar, Abul, 349, 363
Bhirud, Sunil, 615
Biswas, Chitra, 301

C
Cabrera-Alvarez, Elia N., 313
Chakma, Rocky, 263, 301
Chakradhar, Vanam Venkata, 393
Challa, Nagendra Panini, 393
Chaturvedi, Rishabh, 721
Chourasia, Shivani, 289

D
Daniel, Ravuri, 375, 779, 809
Devi, A., 703
Dorado-Sevilla, David F., 61

E
Eck Van, Rene, 145
El Yazidi, Moulay Hafid, 129

F
Fernández-Fernández, Yasmany, 77, 313

G
Gangar, Kavit, 337
Ganiga, Riya R., 557
Gautam, Nishant, 103
Geetha, J., 557
Geetha, M., 569
Ghai, Wiqas, 529
Goswami, Pankaj, 289
Goyal, Nitin, 683
Guleria, Kalpna, 683
Gummadi, Jose Moses, 375

H
Hameed, Mazin Kadhum, 11

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Singapore Pte Ltd. 2021
V. Bindhu et al. (eds.), *International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 733,
https://doi.org/10.1007/978-981-33-4909-4

Hariharan, R., 501
Harikumar, M. E., 175
Harish Annamalai, P., 175
Harmse, Anneke, 157
Herrera-Granda, Erick P., 61
Herrera-Granda, Israel D., 61
Hosatti, Shraddha, 615
Howlader, Moon, 221
Hubackova, Sarka, 255
Hugar, Shobha I., 647

I
Idrees, Ali Kadhum, 11

J
Jayalakshmi, D. S., 557
Jayasekara, K. G. D. R., 25
Jaya Suma, G., 375, 809
Jeong, Seung Ryul, 201
Joshi, Nikhil S., 239
Joy Mathavan, J., 25

K
Kanavalli, Anita, 461
Kannaiya Raja, N., 713
Kannur, Anil, 541
Kanpurrani, V., 653
Kantha Rao, V., 789
Karthika, P., 713
Karthikeyan, S., 501
Kaur, Navneet, 529
Kavya, G., 703
Khadka, Adesh, 201
Kota, Harsha Surya Abhishek, 393
Kottur, Shaguftha Zuveria, 557
Krishna Mohan, A., 375, 809
Kumar, Jitendra, 721
Kumar, N., 579
Kumar, Preetham, 569
Kumar, Yatindra, 693
Kunaraj, A., 25

L
Lakshmi Narayan, B. N., 663, 799
Lakshmi Prasanna, P., 789
Lalthlamuana, James, 673
Latif, Ghazanfar, 87, 349, 363
Laxmi Lydia, E., 375, 779, 809
Lele, Shreyas, 337
Lorente-Leyva, Leandro L., 61, 77, 313

M
Mahadevaswamy, U. B., 239
Mahan, Shriya Raj, 461
Mahbub, Tauhidul, 279
Mahtab, S. S., 263, 279, 301
Maity, Niladri Pratap, 673
Maity, Reshmi, 673
Malatji, William Ratjeana, 145
Mariyan Richard, A., 663, 799
Mazumder, Md. Mahfuzul Islam, 185
Minhajul Alam, S. M., 301
Mishra, Bharat, 821
Mishra, Manish Kumar, 713
Mizan, Mahia Binte, 221
Mohammad, Suhail, 745
Moloto, Makoena, 157
Moses Gummadi, Jose, 779, 809
Mrudula, G. B., 103
Mugunthan, T., 501
Mungurwadi, Vaishali, 647

N
Naina Gowda, K. N., 239
Namrata Mahender, C., 733
Nand, Mekala Pranvitha, 745
Nayak, Sushmitha S., 461

P
Pasha, Syed Nawaz, 383
Patil, Nagaraj B., 43
Pattanaik, Balachandra, 713
Pattanaik, Chinmaya Ranjan, 375
Peddoju, Suresh K., 113
Peluffo-Ordóñez, Diego H., 61, 77, 313
Prabakar, Nagarajan, 113
Pradan, Md. Anuvob, 221
Prasad, B., 779
Prasad, N. H., 663, 799
Prasad, Saurabh, 461
Prasad, Shiv, 461
Priyatham, Manoj, 629

R
Rafie, Marwan, 349
Raihan, Ahamed, 279, 301
Rai, Shwetha, 569
Raja Raman, K., 175
Ramesh, Dadi, 383
Ranjan Pattanaik, Chinmaya, 779, 809
Ranjitha, K. R., 475
Rathipriya, R., 765

Rawat, Sur Singh, 693
Ribin Jones, S. B., 579
Ripon, Shamim, 221
Rolant Gini, J., 175
Ruparel, Hardik, 337, 615

S
Sairamesh, L., 593
Sallauddin, Mohmmad, 383
Samudrala, Sahithi, 745
Sandeep, S., 789
Santhosh Kumar, C., 103
Sarma Dhulipala, V. R., 487
Savakar, Dayanand G., 541
Saxena, Aditi, 721
Sekhar Babu, B., 789
Selvakumar, K., 593
Selvaraj, S., 765
Sharavana Kumar, M. G., 487
Sharma, Amandeep, 683
Sharma, Sudhir Kumar, 289
Sharma, Vivank, 437
Shibly, Mir Moynuddin Ahmed, 185
Shirole, Mahesh, 615
Shrestha, Deepanjal, 201
Shrestha, Deepmala, 201
Shwetha, N., 629
Shyam Mohan, J. S., 393
Singh, Ashutosh Kumar, 821
Singh, L. L. K., 673
Sivakumar, Premkumar, 753
Soni, Jayesh, 113
Soumya Krishnan, M., 419
Sreekumar, K., 449
Sreeman, Vedantham Hanumath, 393
Srinivasa Gupta, N., 437
Sreenivas, Nune, 713
Srivastava, Shobhit, 437
Subramanyam, M. U. M., 393
Sunagar, Pramod, 461
Sundaram, B. Barani, 713
Sunil, G., 383
Surabhi, Tallapalli, 557
Swaroop, Surekuchi Satya, 393
Syed Asif, M. N., 239

T
Telsang, Danesh, 541
Thejaswini, S., 475
Tisha, Tahmina Akter, 185
Tripathi, Satyendra, 821

U
Umaquinga-Criollo, Ana C., 313
Upadhyay, Himanshu, 113
Usha Kumari, CH., 779

V
Vaddi, Radhesyam, 745
Vaijayanthi, S., 515
Valarmathi, B., 437
Vallikannu, A. L., 653
Veerashetty, Sachinkumar, 43
Verma, Sashi Kant, 693
Vhivek, S., 501
Vijayalakshmi, B., 601
Vijay, B., 175
Vimina, E. R., 407, 419, 449
Vinotha, D., 753
Vishva, R., 175

W
Wenan, Tan, 201

Y
Yousfi, Aola, 129

Z
Zellou, Ahmed, 129
Zuva, Keneilwe, 325
Zuva, Tranos, 145, 157, 325